(See Notice on inside front cover)

Vol. 33, No. 6 December, 1943

Journal of Experimental Psychology

EDITED BY

SAMUEL W. FERNBERGER, ON LEAVE
FRANCIS W. IRWIN, ACTING EDITOR
UNIVERSITY OF PENNSYLVANIA

CONSULTING EDITORS

CHARLES W. BRAY, *Princeton University;* ELMER K. CULLER, *University of Rochester;* CLARENCE H. GRAHAM, *Brown University;* JOY P. GUILFORD, *University of Southern California;* FRANCIS W. IRWIN, *University of Pennsylvania;* DONALD G. MARQUIS, *Yale University;* ARTHUR W. MELTON, *University of Missouri;* CARROLL C. PRATT, *Rutgers University.*

CONTENTS

A Further Study on the Bi-directional Goal Gradient in the Endless Maze: M. E. THOMPSON AND C. C. DOVE 447

Studies in Serial Verbal Discrimination Learning: IV. Habit Reversal after Two Degrees of Learning: D. C. MCCLELLAND 457

The Genetic Development of Patterns of Voluntary Activity: R. C. DAVIS ... 471

Experimental Studies of the Judgmental Theory of Feeling: VI. Concrete Versus Abstract Sets in the Preference Judgments of Pictures: H. N. PETERS 487

Problem Solution by Monkeys Following Bilateral Removal of the Prefrontal Areas: IV. Responses to Stimuli Having Multiple Sign Values: H. F. HARLOW AND T. SPAET 500

Accuracy of Recognition of Subliminal Auditory Stimuli: J. W. COYNE, H. E. KING, J. ZUBIN AND C. LANDIS 508

A Simplified Oscillator Suitable for Auditory Research and Audiometry: P. KELLAWAY AND G. BRIGHOUSE 514

PUBLISHED MONTHLY, TWO VOLUMES PER YEAR

BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

PRINCE AND LEMON STS., LANCASTER, PA.

AND NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879

The present issue completes Volume 33. Hereafter, until further notice, this Journal will be published bimonthly, February, April, June, August, October, December, one volume per year, by action of the Board of Editors and the Council of the American Psychological Association. This action has been made necessary by war conditions, which have brought about a decrease in the amount of appropriate material submitted for publication.

Journal of
Experimental Psychology

VOL. 33, No. 6 DECEMBER, 1943

A FURTHER STUDY ON THE BI-DIRECTIONAL GOAL GRADIENT IN THE ENDLESS MAZE [1]

BY MERRELL E. THOMPSON

State University of Iowa

AND

CLAUDE C. DOVE

New Mexico College of Agriculture and Mechanical Arts

I. INTRODUCTION

In 1937 Muenzinger, Dove and Bernstone (5) reported a study with albino rats which presented evidence that the goal gradient is bi-directional. They considered this double winged gradient as a special case of the Thorndike effect (8), i.e., that the goal response 'spreads its effects' both forward and backward. Hull would state this effect in terms of learning being reinforced when the reinforcing state of affairs precedes the S−R conjunction as well as when it follows it.

These results have been confirmed by Hill (2) and Jenkins (4). However, the hypothesis offered by Muenzinger, Dove and Bernstone to account for the bi-directional gradient has not found favor among Hull and his students. Hull (3, p. 162) recognizes the gradient (bi-directional) but dismisses it from theoretical consideration with the following statement: "However, since the reduction in a need necessarily follows rather than precedes the act which brings this about, it would seem that the forward gradient could hardly play much role in selective learning." [2]

[1] This study was made at the New Mexico State College of Agriculture and Mechanical Arts and submitted by M. E. T. in partial fulfillment for the degree of Master of Science.

[2] One might as well inquire as to the role played in selective learning by Hull's primary gradient of reinforcement which falls to zero around 30 sec. (3). See also Perin (6, 7).

It is easy to overlook the fact that the ordinary maze situation is highly artificial in nature and that in life situations organisms continually move from one goal to another.

Hill (2) points to several of his findings which he claims to invalidate Muenzinger. Dove and Bernstone's hypothesis concerning the forward spread of reinforcement. On p. 578 he says, "This view seems untenable in this experiment for several reasons. First, the single gradient is present early in training. There seems to be no reason why the double gradient should not be existent early in training if it is the result of this forward spread of reinforcement." This point is well taken, and if we discover in the true goal-to-goal situation a single (backward) gradient existing early in the learning series it would constitute a serious blow to the hypothesis of Muenzinger, Dove and Bernstone. It is unfortunate that Muenzinger, Dove and Bernstone failed to present data in their study bearing on this factor.[3]

Hill's failure to demonstrate the presence of this double gradient early in training is perhaps due to his experimental procedure, which was not precisely a goal-to-goal situation. Certainly Hill's method of picking the animal up from the food box and returning him to the starting compartment after each daily run is not directly analogous to the procedure followed by Muenzinger, Dove and Bernstone.[4] Hill also failed to run a control group and neglected mentioning the start-to-goal group reported by the above mentioned authors.

Hill stated his strongest point as follows: "probably the most convincing evidence against the view that the bi-directional gradient observed here is a function of bi-directional reinforcement is the data from Table III, which shows that the bi-directional gradient arises from a failure to eliminate anticipatory errors adequately at the middle choice points" (p. 578). An analysis of this statement does not appear to clarify the issue. There is no question of the middle errors not being eliminated as rapidly as errors at the beginning and last points of choice—this is obvious from the nature of the bi-directional curve itself. It is true that Hill's analysis does show that these errors are anticipatory ones. However, he fails to show the mechanism underlying the phenomena. The important point, as the writers see it, is: *why* do the middle errors eliminate more slowly than the errors at the beginning and end of the maze?

[3] The present study was carried out prior to Hill's publication. However, an analysis was made on the development of the bi-directional gradient and presented in 1938, in thesis form.

[4] It is presumptuous to assume that in Hill's study the animals were running from goal-to-goal. Hill makes his comparison on the basis of Brandt's study (1) which showed that the spread of reward from a successful response is not a function of the time between units but of the number of units. To assume, as Hill did, that the fact that the animals were returned to the starting compartment immediately after consuming the food in the goal box constituted a goal-to-goal sequence is omitting consideration of certain other important factors. We can not assume that what happens to the animal *outside* the maze has no significance. Such things as picking up, handling, moving the animal through space, etc., would seem to play an important role in the animal's subsequent behavior in the first part of the maze.

II. Purpose of the Study

The purpose of this study was to examine further the nature of the bi-directional goal gradient and to check the findings of Muenzinger, Dove and Bernstone (**5**). A new maze was constructed which eliminated the weakness of the four-section square maze used by them. This new maze was a double-straightaway maze consisting of two identical sections with seven hurdles of seven doors each. This made possible a more successful patterning of open doors as the animal ran through a different door at each hurdle while running any particular pattern. (See Fig. 1.)

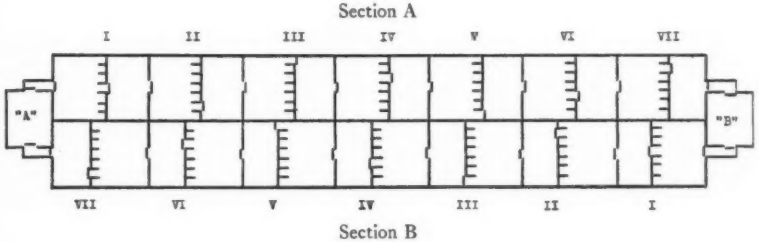


Fig. 1. The double-straightaway maze used in the experiments on the goal gradient. In Experiment II, the start-to-goal group, the animals ran through section 'A' only.

The Roman numerals indicate the primary partitions or hurdles at which the animals had to learn the correct (open) doors. The other, secondary, partitions contained a center door which was always unlocked.

The open or unlocked doors shown in the figure indicate pattern No. VI of the patterns arranged for these experiments.

The two food boxes are marked 'A' and 'B.'

III. General Experimental Procedure

The entire experiment was carried out with 87 albino rats of Wistar strain between 90 and 150 days old at the beginning of the experiment. The apparatus employed for the experiment was an endless maze of the double-straightaway type in which the animals had to run from goal-to-goal, i.e., from food box to food box. The maze consisted of two straight sections 14 ft. long, 18 in. wide, and 5½ in. high. It was constructed of wood and covered with removable screen wire. The entire maze was painted a dull black to reduce visual cues as much as possible. Constant illumination was maintained by suspending individual lights over each choice point. The seven choice points in each section of the maze consisted of seven identical doors of the single-valve type. Each door was faced by a short open top alley pointing in the direction of the food box. There were also other partitions with a single door located midway between the choice points. This single door was located in the center of the partition and remained unlocked throughout the experiment. These secondary partitions with the center door served to force the animal to make his choice from the center of the maze at each choice point and also prevented retracing.

The study consisted of two experiments. The first, or goal-to-goal experiment, employed 36 animals divided into seven groups for patterning purposes. Experiment II was made with 51 animals run in a start-to-goal situation.

In the preliminary training period all animals, in both groups, were taught to go under the doors. They were trained in a straight maze patterned after the experimental maze only on a smaller scale. The doors in this small maze were always unlocked and each animal was trained

to run through the doors to a small piece of cheese. After learning to open the doors the animal was placed in the experimental maze and allowed to make one run with all doors unlocked. In the goal-to-goal group the animal was placed in the starting compartment where he found a bit of cheese. Then, after running through the seven choice points, he found another small piece of cheese in the goal box. Animals in the start-to-goal group received the same kind of training, but they were not fed at the beginning of the run. In both groups the experimental runs were not started until the day following this preliminary training.

When the experimental series began each animal in the goal-to-goal group was fed a bit of cheese in the goal box from which it began the daily runs and was rewarded with similar bits of cheese in each of the goal boxes thereafter on its way around the endless maze. All animals began the daily runs from the goal box marked 'A' in the diagram (Fig. 1). Each animal ran around the maze five times each day for the first five days, i.e., it ran through the same maze pattern 50 times during the first five days. After this the animal was given seven or eight runs around the maze each day until he had made a total of 50 complete runs, or 100 runs on the same pattern. Many of the animals had learned the maze before the end of this procedure but were given the daily runs until they had made 100 runs.

The apparatus was slightly changed for the start-to-goal group. Only one section (A-Section) of the maze was used. The food box ('A') was replaced with an identical box in which no food had been placed. An animal was placed in the starting box without food and ran through this one section of the maze to food and was then lifted out of the goal box and returned to the home cage. (Animals in both experiments were fed a balanced diet in the home cage immediately following the daily runs.) On the following day the procedure was repeated until each animal had made 50 runs. Thus an animal in the start-to-goal group was given one trial per day for 50 consecutive days.

EXPERIMENT I

Goal-to-Goal Group

Purpose: The purpose of this experiment was to check the results reported by Muenzinger, Dove and Bernstone (**5**) using an improved maze situation.

Apparatus and Animals: The apparatus used in this study was the double-straightaway endless maze (Fig. 1). Thirty-six albino rats were used as subjects.

Procedure: For experimental purposes the 36 rats were divided into seven groups. This grouping was determined largely by the time of the experimental work and the number of animals available at any time. The groupings were as follows:

Group	Number of Animals	Pattern
I	6	3-1-6-4-7-2-5
II	6	1-6-4-2-5-7-3
III	6	5-3-1-6-2-4-7
IV	6	2-7-5-3-6-1-4
V	4	7-5-3-1-4-6-2
VI	4	4-2-7-5-1-3-6
VII	4	6-4-2-7-3-5-1

The pattern for the different groups was arranged before the experiment began. See Fig. 1 for a ground-plan of section 'A' and 'B' of the maze and an illustration of a pattern of correct or unlocked doors. The doors in the various primary partitions were numbered from 1 to 7, beginning with the inside door each time. The numbers in the pattern indicates the doors that were open in each partition beginning with the first choice point. Each of the two sections of the maze had the same pattern for any particular group of animals. After running through the seven choice points in one section and eating, the animal was confronted with the same pattern of correct doors in the second section.

In this experiment the animals had to learn to run from goal-to-goal, i.e., from food box to food box. Only one door at each choice point was open during the experimental series and they were open in accordance with the above patterns.

Results: The composite results of the seven groups in Experiment I are given in Table I below and are shown graphically in Fig. 2.

TABLE I

COMPOSITE RESULTS OF EXPERIMENT I

Hurdle	Section 'A'		Section 'B'		Composite	
	Errors	Average	Errors	Average	Errors	Average
I	466	12.94	377	10.47	843	23.42
II	543	15.08	453	12.58	1003	27.86
III	678	18.84	539	14.97	1217	33.80
IV	802	22.27	620	17.22	1422	39.50
V	666	18.50	549	15.25	1214	33.72
VI	509	14.13	430	11.94	939	26.08
VII	362	10.05	281	7.80	643	17.86

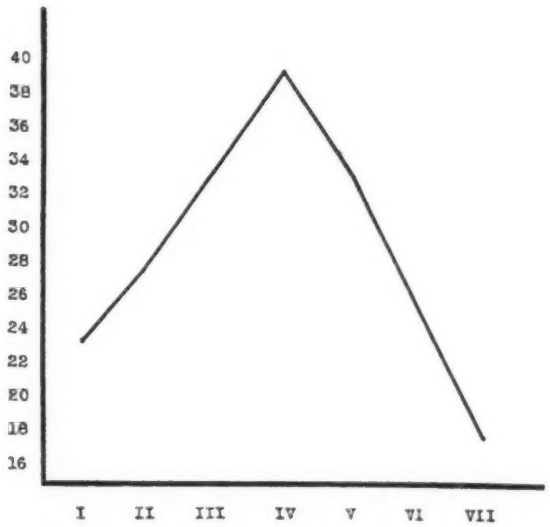


FIG. 2. Mean number of errors made by the 36 rats in the goal-to-goal group. A total of 100 runs.

The critical ratios for section 'B' for different hurdles in the goal-to-goal group are as follows:

Between hurdle 1 and 4 ..... 3.7
Between hurdle 1 and 7 ..... 2.3
Between hurdle 4 and 7 ..... 5.7

The results confirm the bi-directional gradient found by Muenzinger, Dove and Bernstone. There was no break in the gradient between the middle choice point and the one following as occurred

in Muenzinger, Dove and Bernstone's gradient. The maze and pattern correction perhaps account for this.

EXPERIMENT II

Start-to-Goal Group

Purpose: The purpose of this experiment was to check our maze and general experimental set-up for the old typical uni-directional goal gradient.

Apparatus and Animals: The apparatus used in this study was one section of the double-straightaway maze shown in Fig. 1. Fifty-one albino rats served as subjects.

Procedure: The 51 rats were divided into seven experimental groups. The groupings and patterns run were as follows:

Group	Number of Animals	Pattern
I	7	3-1-6-4-7-2-5
II	8	1-6-4-2-5-7-3
III	8	5-3-1-6-2-4-7
IV	6	2-7-5-3-6-1-4
V	7	7-5-3-1-4-6-2
VI	5	4-2-7-3-1-3-6
VII	10	6-4-2-7-3-5-1

In this experiment the animals ran from start-to-goal. The procedure was identical to that employed in Experiment I with two exceptions: (1) the animals received food at the end of the maze only; and (2) the learning series was done under distributed practice, i.e., one run per day.

Results: The composite results of the seven groups in Experiment II are given in Table II below and are shown graphically in Fig. 3.

TABLE II

COMPOSITE RESULTS OF EXPERIMENT II

Hurdle	Total Errors	Average Errors Per Rat
I	1158	22.70
II	962	18.86
III	877	17.19
IV	975	19.08
V	884	17.33
VI	858	16.82
VII	658	12.90

The critical ratios for different hurdles in the start-to-goal group are as follows:

Between hurdle 1 and 4	3.5
Between hurdle 1 and 7	10.3
Between hurdle 4 and 7	9.9

The results of Experiment II confirm the conclusions drawn by numerous studies in animal psychology on the nature of the orthodox goal gradient. Here, as in other reported studies, we have the backward elimination of errors from the goal (food box) when the animals run in a start-to-goal situation.

IV. Development of the Bi-Directional Gradient

If the double gradient is to be explained as a function of the forward-backward spread of reinforcement from the goal, it is important to show that it develops as the maze is learned, i.e., that it is a function of learning itself. We have seen that Hill criticized this explanation, since he found a single gradient early in the learning, the double gradient making its appearance only in the later trials.

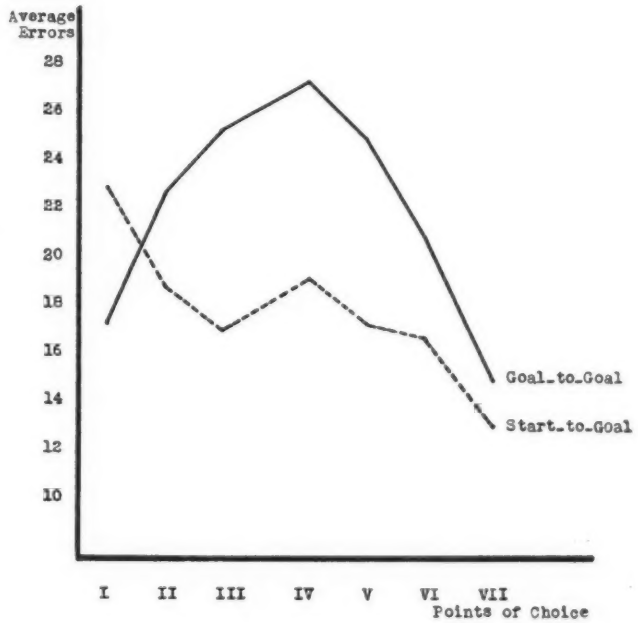


Fig. 3. Error Gradients

The broken line represents the average errors made by the start-to-goal group, the solid line that of the goal-to-goal group (first 50 trials).

In the first experiment reported in this study we find a pronounced double gradient on the tenth run. This bi-directional gradient becomes more pronounced as the maze is learned, showing the maximum spread on the twentieth run. An examination of Fig. 4 will show that the animals had practically learned the first and last hurdle perfectly by the 20th run, the flattening out of the gradient in later stages of the series illustrates an approach to perfection in running the maze.

Listed below are the total errors made on the 10th, 20th, 30th, 40th and 50th run made in section "B" by the 36 animals. This is shown graphically in Fig. 4.

On Choice Points	I	II	III	IV	V	VI	VII
Total errors on 10th run	14	18	19	23	21	11	13
Total errors on 20th run	4	5	11	17	10	5	3
Total errors on 30th run	3	3	7	7	7	2	2
Total errors on 40th run	0	3	6	5	4	0	3
Total errors on 50th run	1	2	3	4	1	0	1

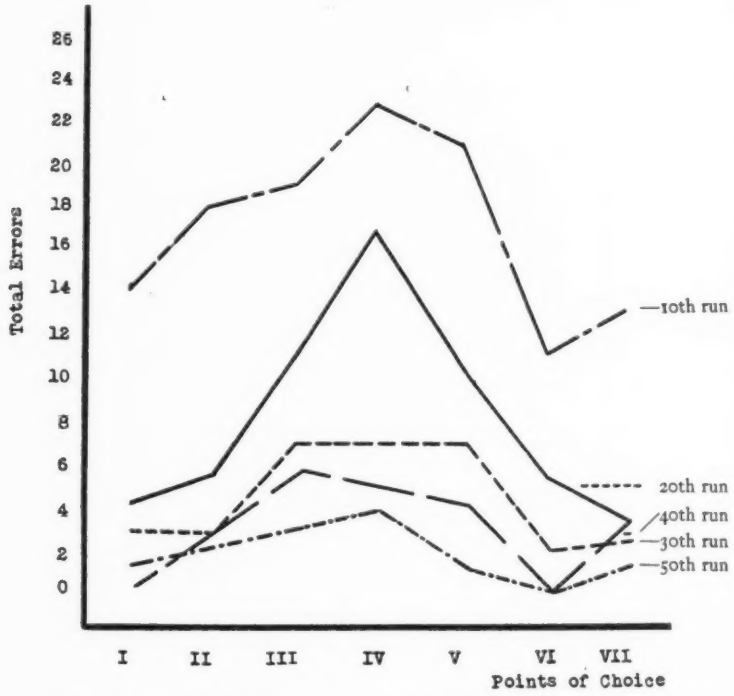


FIG. 4. Error gradient showing the development of the bi-directional gradient. This shows total errors made by the 36 rats in the goal-to-goal group.

V. DISCUSSION

It has been shown that a double gradient exists from the early learning series of the goal-to-goal group. This would be expected if the hypothesis presented by Muenzinger, Dove and Bernstone were correct. The appearance of a double gradient later in Hill's study

may be due to an emotional adaptation of his animals to the break in running the maze. This would explain the discrepancy between the two gradients found in the early and late stages of his learning series. In the early stages of his learning the animals were not running from goal-to-goal due to the change in sequence brought about by being removed from the maze and replaced in the starting box for the next run. However, as this procedure continued the animals would become accustomed to the handling, thus giving the reinforcement a chance to affect the choice points in the first part of the maze. This explanation is not presented as a rigorous deduction of Hill's results; it is suggested as a possible means of reconciling his results with ours.

Before final acceptance of the forward-backward spread of reinforcement hypothesis, more empirical investigations must be made. One variable which has not been controlled in these studies is that of massed and distributed practice. Is it possible to demonstrate that the uni-directional gradient becomes a bi-directional one under massed practice? Hill's study certainly indicates as much. If this is the case, can we still explain both gradients by Hull's goal gradient hypothesis or would the hypothesis presented by Muenzinger, Dove and Bernstone prove more adequate? At the present time Hull has made no attempt to incorporate this extension of the goal gradient into his theory.

The scope of the present paper does not permit a presentation of a rigorous quantified hypothesis concerning the bi-directional goal gradient. However, several pertinent factors should be discussed in order to throw some light on the problem. It is necessary to examine Hill's results and comments further if we are to gain a clear picture of the issue. Hill offers an ingenious and penetrating analysis of his data which appears to eliminate the possibility of a forward spread of reinforcement. However, this minute description of the nature of the errors which cause the appearance of a double-winged gradient is meaningless as far as explanation is concerned. It seems, to the writers at least, that labeling the errors anticipatory does not invalidate the original hypothesis concerning the forward spread of reinforcement in the least.

The theory offered here may be stated briefly as follows:

(1) Food in a goal box gains secondary reinforcing properties because of its ultimate need reduction. This secondary reinforcement is active both forward and backward.

(2) Various cues in the choice points preceding the goal gain secondary reinforcing properties in relation to their distance from the goal. Thus the choice point just preceding the goal will acquire more secondary reinforcing properties and be learned before the choice point preceding it.

(3) The properties of the choice points immediately following the goal will also acquire secondary reinforcing properties proportional to their distance from the goal. A trace of certain of the components involved in the eating and swallowing of food will persist for a short time as traces. It seems logical that cues regularly associated with the taste trace, etc., will gain power of reinforcement and this reinforcement is perhaps as great as that gained in the hurdles preceding the goal.

VI. Summary

In a maze with seven seven-door equidistant choice points, 87 albino rats were run in two experimental situations. Thirty-six animals were given from 10 to 16 runs a day in a goal-to-goal situation. Fifty-one animals were given one run a day for 50 consecutive days in a start-to-goal situation.

1. The record made by the goal-to-goal group showed a forward-backward elimination of errors, i.e., the first and last correct choices were learned before the other five choice points.

2. This bi-directional gradient was shown to appear in the early stages of learning and reach its maximum spread on the 20th run.

3. The record made by the start-to-goal group showed the typical backward elimination of errors.

4. It was shown how Hill's results could be deduced without invalidating the hypothesis presented by Muenzinger, Dove and Bernstone.

5. A brief theory was presented which explains the double-winged gradient in terms of the forward-backward spread of secondary reinforcement.

(Manuscript received November 10, 1943)

Bibliography

1. Brandt, H. The spread of the influence of reward to bonds remote in sequence and time. *Arch. Psychol.*, N. Y., 1935, **27**, No. 180.
2. Hill, C. J. Goal gradient, anticipation, and perseveration in compound trial-and-error learning. *J. exper. Psychol.*, 1939, **25**, 566–585.
3. Hull, C. L. *Principles of behavior.* N. Y., 1943, D-Appleton Century.
4. Jenkins, W. O. Studies on the spread of effect. I. The bi-directional gradient in the performance of white rats on a linear maze. *J. comp. Psychol.*, 1943, **35**, 41–56.
5. Muenzinger, K. F., Dove, C. C., & Bernstone, R. H. Serial learning: II. The bi-directional goal gradient in the endless maze. *J. genet. Psychol.*, 1937, **50**, 229–241.
6. Perin, C. T. A quantitative investigation of the delay-of-reinforcement gradient. *J. exper. Psychol.*, 1943, **32**, 37–51.
7. Perin, C. T. The effect of delayed reinforcement upon the differentiation of bar responses in white rats. *J. exper. Psychol.*, 1943, **32**, 95–109.
8. Thorndike, E. L. An experimental study of rewards. *Columbia University Contributions to Education*, 1933, No. 580.

STUDIES IN SERIAL VERBAL DISCRIMINATION LEARNING. IV. HABIT REVERSAL AFTER TWO DEGREES OF LEARNING[1]

BY DAVID C. McCLELLAND

Wesleyan University

There have been a number of theoretical treatments of the mechanism of discrimination learning (**1**, **2**, **7**). But, as Hilgard and Marquis point out (**5**, p. 192), though the treatments differ in detail, they all make certain common assumptions. One of the most important of these is that the choice reaction which will be made at any given moment in learning is that one which "has the greater excess of positive reaction tendencies conditioned to it" (**5**, p. 193). That is, given two reaction tendencies—to go right or to go left—that one which is strongest, which has received the most reinforcement in the past, is the one which will be made. It is usually further assumed that the greater the advantage in strength that one reaction tendency has over another, the greater is the certainty that the stronger will be made.

A deduction which follows directly from these assumptions is that reversal of a discrimination will become increasingly difficult as original training proceeds. That is, increasing practice on one response in a discrimination transfers increasingly negatively to the reverse response. Confirmation of the deduction in animals is to be found in the recent work of Williams (**19**) and in earlier studies (**9**, **12**). Spence in particular has emphasized the theoretical importance of the deduction and has stated it explicitly (**17**, p. 448) in his controversy with Krechevsky (*cf.* **8**, **18**) over the continuity theory of discrimination learning.

At the human level the deduction appears to be an oversimplification. For one thing studies with negative practice (*cf.* **15**) have shown that practice in spelling 'hte' (i.e., strengthening one response) may transfer positively to spelling another response 'the.' It may be, however, that this does not make an exception to the reversal deduction, since the *S* practices with knowledge that he is wrong, a condition which may serve to weaken rather than strengthen the wrong response.

[1] This study was carried out in the psychological laboratory at Connecticut College. I should like to acknowledge also the generous help and criticism provided by Dr. K. S. Lashley and Dr. Kenneth Spence in the initial stages and by Dr. C. J. Hill in the final preparation of the report.

A more convincing case against the reversal deduction may be made on the basis of common-sense introspection. From this point of view the *S* should have the greatest difficulty in making a correct reaction when the two alternatives are equally strong. As soon as either becomes stronger than the other he should be able to choose *either* more or less consistently. This means that reversal of a discrimination should become easier as original training proceeds, in apparent contradiction to the usual reversal deduction.

The present experiment was designed to throw some light on this apparent contradiction by creating a situation in which the transfer from practice on one response of a discrimination to the other could be measured in human *S*s. This situation differs from the usual negative practice experiment in that the original training is done without the knowledge that the response is 'wrong' relative to a response required at some future date. Consequently the results will also show whether there can be positive transfer from negative practice which is not known to be wrong.

The essential parts of the experimental design were (*a*) to require the *S*s to learn one of two responses to a series of words, (*b*) to ask them to make the opposite response to each word after two degrees of training on the original reactions, and (*c*) to note by comparison with an unreversed group how much they had gained or lost from each of the two degrees of original training.

PROCEDURE

The experimental method used was essentially the same as that reported in a previous study of serial verbal discrimination learning (11). A list of 15 adjectives was prepared and presented to the *S* at the rate of one every two sec. in three different random orders on an ordinary Hull-type memory drum. The list was shortened from the 20 cue stimuli previously used to 15 because it was thought that reversal might make relearning impossible or extremely tedious if the list were so long. The *S*'s task was to learn to say aloud certain words in the list and not to say aloud certain others. He began by guessing on the first trial, saying some of the words aloud and not saying others. If he happened to make the correct response to a word (which involved either speaking or not speaking depending on the word), he was 'rewarded' by a short ring from a bell, the circuit of which was run through a portion of the memory drum so that if the experimenter closed a key, it would ring just as the drum turned to the next word. The instructions used with this technique have been described in full elsewhere.

TABLE I

Condition	Original Learning	Rest	Relearning 1	Rest	Relearning 2
S (control).....	Pattern A to 11/15	30 sec.	Pattern A to 15/15	30 sec.	Pattern A to 15/15
R'.....	Pattern A to 11/15	30 sec.	Pattern B to 15/15		
R.....	Pattern A to 11/15	30 sec.	Pattern A to 15/15	30 sec.	Pattern B to 15/15

Pattern A is the reverse of Pattern B, that is, all the words which should be spoken aloud in Pattern A should not be said in Pattern B and vice versa. Some *S*s began with Pattern A and some with Pattern B. *Cf.* Table II. The symbols 11/15 and 15/15 refer to criteria of learning, 11 and 15 out of 15 possible correct choices respectively.

The *Ss* were 24 girls, undergraduates at Connecticut College. Each served on four days, the first for practice. On this day they became thoroughly familiar with the recitation technique and learned one list to the mastery criterion of one perfect trial. When they had reached a criterion of 11 in 15 or 15 in 15 correct responses in the learning of the practice list, they were asked to reverse their responses to the words, that is, to make the opposite and formerly wrong response to each word. This was done to familiarize them with the reversal procedure so that when it was called for on the work days they would adopt it with a minimum of disruption.

The three work conditions are outlined in Table I. In all three the *Ss* learned to a criterion of 11 out of 15 possible correct responses and then rested 30 sec. This did not represent a high degree of learning since 7 or 8 correct choices would be made by chance alone. In condition S (same response to each stimulus required throughout), the *Ss* continued learning the same responses after the rest to one perfect trial, then rested 30 sec., and relearned to a second perfect trial. In condition R′ (reverse response to each stimulus required after partial learning) the *Ss* were instructed during the 30-sec. rest to reverse their responses to each stimulus word on the next trial. They then continued learning the reversed responses to one perfect trial. In condition R (reverse response to each stimulus required after complete learning) after the 30-sec. rest the *Ss* continued learning the same responses as in condition S to one perfect trial. Following this, during a 30-sec. rest they were instructed to reverse their responses and were practiced until they had learned the reverse responses to one perfect trial. The experiment was designed in this fashion so that the control group would serve for comparison with each experimental group. Relearning scores would be obtained for reversed and unreversed responses after two degrees of original learning.

Complete counterbalancing procedures were followed out. The variables involved were: Lists 1, 2, and 3; conditions S, R′, and R; Patterns A and B; and degree of learning on the practice day at which reversal was introduced (11/15 or 15/15). Table II indicates how these factors were varied systematically for the first six *Ss* so as to counteract any interaction effects.

TABLE II

COUNTERBALANCING PROCEDURE FOR SIX *Ss*

	Successive Days		
Practice	Work		
1	2	3	4
11/15 4A	S 1B	R 2A	R′ 3B
15/15 4B	S 1A	R′ 3B	R 2A
15/15 4A	R 3B	R′ 1A	S 2B
11/15 4B	R 3A	S 2B	R′ 1A
15/15 4B	R′ 2A	R 1B	S 3A
11/15 4A	R′ 2B	S 3A	R 1B

The symbols 11/15 and 15/15 refer to the criterion reached on the practice list when the *Ss* were asked to reverse; 1, 2, 3, 4 to the list used; A and B to reciprocal patterns of responding; S, R′, and R to the experimental condition.

Since complete counterbalancing was impossible with six subjects, three more similar procedures were worked out for the three other groups of subjects such that each condition appeared with each list, each list with each pattern, each pattern with each condition, each condition with each day, etc., all an equal number of times.

RESULTS

Before comparisons among the conditions can be made it must be shown that they were equated in original learning. The average number of correct responses on the criterial trial was 11.42, 11.58,

and 11.54 for the S, R′, and R groups, respectively, indicating a satisfactory equality in this measure. The average number of trials taken to reach the criterion of 11/15 was 8.50, 6.46, 6.96 for S, R′, and R groups, indicating unsatisfactory equality. The control group S took 2.04 trials longer on the average than the work group R′ with which it was to be compared, a difference which is 3.18 times its standard error ($P < .01$) and hence extra-chance.[2] Logically, however, there is no way that this difference could have arisen except by chance. The complete counterbalancing procedure absolutely precluded any differential treatment of the *S* under any of the three conditions. A thorough analysis of the possibilities that the difference was due to list effects, practice effects, or interaction effects among conditions showed that none of these factors could explain it. The only remaining explanations are either that it *was* chance despite the improbability or that it was due to some complex interaction effects of the three factors.

TABLE III

COMPARISON OF MEAN SCORES TO LEARN SAME AND REVERSE RESPONSES TO WORDS IN A LIST AFTER ORIGINAL LEARNING TO A CRITERION OF 11 IN 15 CORRECT RESPONSES. N = 24.

Condition	No. Correct, Criterial Trial	No. Correct, 1st Post-criterial Trial	No. Correct, 2nd Post-criterial Trial	Trials to One Perfect
R (same)	11.54	11.08	11.17	9.79
R′ (reverse)	11.58	10.08	9.63	13.00
Differences (R−R′)		1.00	1.54	−3.21
$\sigma_{diff.}$		.53	.62	1.68
Critical ratio		1.89	2.49	1.91
P		<.05	<.02	<.05

P is the probability that as large a difference *in the obtained direction* could have occurred by chance in samplings of populations of this size.

Fortunately, because of the design of the experiment, the R group can serve as a control for the R′ group. The mean trials to reach the first criterion of 11/15 (6.96 and 6.46) do not differ significantly. And the S group may still serve as a control for the R group after the second criterion (15/15) is reached, since the mean trials taken by the two groups to reach this point (18.58 and 16.75 respectively) do not differ significantly (C.R. = 1.24, $P < .13$). Because of these considerations, subsequent comparisons will be made after partial

[2] The standard errors of differences in this experiment are all computed by the direct method so that the correlation between scores made by the same *S* may be automatically taken into account. In subsequent statements of the reliability of differences C.R. stands for critical ratio, P for the probability that such a difference in the obtained direction could have occurred by chance.

learning between the R and R′ groups and after complete learning between the S and R groups.

Table III presents for comparison the scores obtained for relearning the same responses and for learning the opposite responses after the criterion of 11/15 correct responses had been reached. Since the reversed group is consistently poorer on the first and second post-criterial trials, and since this group takes reliably longer to attain one perfect trial, the conclusion appears amply justified that requiring the reverse response to each stimulus after partial learning slows learning over what it would be without reversal.

But the significant fact is that the decrement is not large. On the contrary, if the reversed learning scores are compared with learning without any previous practice, they show large positive transfer. The average number of correct responses on the first trial, 10.08, is significantly larger than the average number after no previous practice, 7.38 (C.R. = 4.45, $P < .01$). Similarly, the number of trials taken to reach one perfect trial, 13.00, is 3.75 trials less on the average than the number taken without previous practice, 16.75, a difference which is significant (C.R. = 2.21, $P < .03$) and which represents a saving of 23.81 percent.[3] The saving in trials to reach the criterion of 14/15 is even greater (37.73 percent), the difference larger (4.92) and more reliable (C.R. = 3.91, $P < .01$).

TABLE IV

Comparison with Other Similar Means of Mean Scores to Learn Reversed After Original Learning to 15/15

	Mean Correct Responses on First Trial	Mean Trials to One Perfect Recitation
Reversal after OL to 15/15	14.08	3.50
Non-reversal after OL to 15/15	14.33	1.83
Difference	−.25	−1.67
$\sigma_{diff.}$	.23	.52
Critical ratio	1.09	3.21
P	<.18	<.01
Reversal after OL to 15/15	14.08	3.50
'Reversal' after no previous learning	7.38	16.75
Difference	6.70	13.25
$\sigma_{diff.}$	.47	1.00
Critical ratio	14.28	13.25
P	<.01	<.01
Reversal after OL to 15/15	14.08	3.50
Reversal after OL to 11/15	10.08	13.00
Difference	4.00	9.50
σ_{diff}	.50	1.32
Critical ratio	8.00	7.20
P	<.01	<.01

[3] Savings or transfer percents are computed by the formula: (Control — Work)/Control, after the criterial trial has been subtracted from each score.

In view of these figures there can be no doubt that practice on the reverse responses to each of the stimuli up to the criterion of 11/15 is far better than no practice at all. To put it another way, partial training on a series of discriminations transfers positively to learning the reverse discriminations.

What is the effect of increasing original training before reversal? Animal studies (**19**) would lead one to expect that reversal should be more difficult, that the positive transfer thus far obtained, which might be attributed to a general practice or familiarization effect, should decrease or at least grow no greater. Table IV supplies the data which answer the question. In this table the average number of correct responses on the first reversal trial after original learning to a criterion of 15/15 (14.08 responses) is compared: (1) with the average number of correct responses on the similar trial without reversal. The reversal causes a small and unreliable loss. (2) With the average score on the first trial without previous training of any sort. The high degree of training on the reverse responses is comparatively very beneficial. (3) With the average score on the first reversal trial after original learning to the lower criterion of 11/15. The added training on the reverse responses produces a large and significant gain in the number of correct responses on the first reversal trial. To sum up: the more the training on the original responses the greater the transfer to the reverse responses to the same words. There is not only no increasing difficulty of reversal, as one might expect, but actually increasing ease.

The same effect is shown in the second column of Table IV when the scores are trials to reach the criterion of one perfect recitation. According to this measure, requiring reversal after original training to 15/15 reliably increases the number of trials taken to reach one perfect trial as compared with a non-reversed group. But reversal requires 13.25 fewer trials on the average than if no previous training had been given and 9.50 fewer trials than if the original training had proceeded only to 11/15. Unquestionably, again, learning one set of responses to a series of words transfers almost completely to learning the reverse set.

Fig. 1 presents graphically in terms of savings percents the effect of the two degrees of original training on subsequent reversed and non-reversed learning. Each point on the two upper curves represents the percent of the trials needed normally to reach one perfect trial saved by giving two degrees of reversed and unreversed learning. The curve representing percent saved by the group which had reversed training follows very closely the group which had unreversed training, indicating very little disadvantage. It also goes in the opposite direction from the line labelled 'Deduced Reversed' which

represents the results which would be expected from the principle based on animal studies that difficulty of reversal increases with amount of original training.

There are two further sources of evidence that the animal learning principle cannot apply to these data. According to the theory behind the principle, errors which the *S* made in original training, since by definition they are stronger than the right responses, ought to become right responses after reversal. This is not what happens here. In R′ on the last trial of original learning the *S*s made a total of 82 errors. Of these, 45, or 54.88 percent, were still classified as

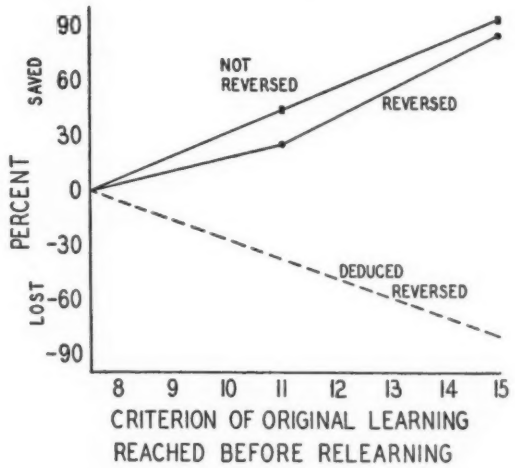


FIG. 1. The percentages saved in relearning reversed and not reversed discriminations after two degrees of original learning compared with the percent losses to be expected from the reversal experiments in animal discrimination learning.

errors on the next trial after reversal, when they would have been counted as successes had the same responses been repeated. In other words, the *S*s *reversed the responses they had been getting wrong* as well as the ones they had been getting right. This figure compares with 49 out of 83 errors repeated (59.04 percent) in the non-reversal group, a difference which is insignificant and which emphasizes the similarity of the reversed and non-reversed groups.

Furthermore, according to animal reversal data the maximum difficulty should occur immediately after reversal. There is some evidence from the R′ group that the maximum difficulty occurs later on. In Table III, for instance, the mean loss caused by reversal is less on the first post-criterial trial than on the second, though the difference is not significant. Furthermore, individual *S*s frequently

did well on the first few reversal trials only to show points of greater confusion later on. It proved impossible to get composite curves which would show this effect because the point of maximum difficulty occurred at different places in relearning for each *S* and combining averaged out the effect. Some evidence on the point was obtained by determining for the R and R′ groups the percent of the total number of correct responses made in each tenth of the relearning period according to the procedure outlined by Hilgard and Campbell (3). If the greatest difficulty in learning reverse discriminations occurred immediately, one would expect that more *S*s would show the lowest percent values in the first tenth of relearning. Actually, nine *S*s showed the lowest percent correct responses in this tenth, and 15 in some later tenth for the reversed group, a proportion which is exactly the same as in the non-reversed group and which emphasizes once more the similarity of the two groups. There is no greater tendency for a reversed group to show its poorest scores early in relearning than for a non-reversed group. The difficulty encountered by the reversed group may be greater, but its locus is distributed over relearning just as it is for a non-reversed group.

Discussion

The data from this experiment do not support the conclusion based on animal studies that reversal of a discrimination will become increasingly difficult as degree of original training increases. On the contrary, training on one set of responses in a number of discriminations transfers positively to learning the reverse responses, and the amount of positive transfer increases proportionately with the increase in original training. It might be argued that the initial positive transfer was due to general practice factors such as familiarity with the technique, the rapid responding required, etc. That is, the group which has had the previous training, even though reversed, might have a certain advantage over a group starting out to learn *de novo*. This argument would be more convincing if both groups had not had considerable practice the previous session. It also seems questionable in view of the fact that the transfer is so large and the fact that it increases so much with increased reversal training. Such general practice might cause the middle curve in Fig. 1 to be shifted upwards throughout, i.e., it might have exaggerated the positive transfer; but it cannot alter the main conclusion that ease of learning reversed discriminations increases, rather than decreases, with the amount of original training.

Such a conclusion is not only out of line with the animal results, but it also does not fit in with the usual theory as to the conditions which cause negative transfer. According to this theory the condi-

tion which produces negative transfer is requiring a new response to a stimulus to which another response has already been attached (**13**, p. 416; **20**, p. 202). Supposedly the stimulus tends to evoke the old response as the new one is being learned, the two responses conflict, and learning is slowed over what it would be without previous training. This is what should happen here since the conditions conform to the negative transfer paradigm. The S−R connection, A−K, is required after the connection A−B has been established. The unique fact, the fact which makes the situation different from those usually producing negative transfer, is that responses B and K are reciprocal. They are the only two alternatives and one of the two must be made.

This fact provides the key to an explanation of the results which will serve to bring them into line with the animal studies and other negative transfer phenomena. It means that in rejecting one response to a stimulus word the *S* practically automatically made the reverse response. If he succeeded in inhibiting through his continuing reversal set all the responses he learned originally, he must have made the correct responses. That is, the already learned responses could have been used as cues to making the newly-required reverse responses. Suppose, for example, he saw a certain word and remembered that he had been learning to say it. If he succeeded, through his reversal set, in checking this response before it became vocalized, he would already have made the correct, reverse response of keeping quiet. Or, similarly, if he remembered that he had been learning not to say a certain word, he could check that response and make the only alternative one of saying it. Consequently, the amount of positive transfer to reversed learning should be a direct function of the amount of original training, just as it is, since the ability to reverse depends on the certainty with which the formerly correct responses are evoked. The positive transfer is not complete (is less than from unreversed learning) because reversal takes time and because it depends on the ability of the *S* to check his formerly acquired responses.

Such an explanation means that these results do not contradict the reversal deduction made by Spence (**17**, p. 448) from his theory of discrimination learning in animals. The data do require that the theory be extended, if it is to apply to human discrimination learning, to include certain mechanisms necessary to the above explanation. The most important of these mechanisms are: a reversal set and symbolic or anticipatory representation of responses. If rats possessed these two abilities it is likely that reversal with them would result in the positive transfer obtained here.

Rats probably can make anticipatory responses. Hull (**17**) has

assumed they can. So the chief difference between this and the animal studies showing opposite results (19) is that the rats did not know, as the human Ss did, that the reverse response was required at a given point in the training. Furthermore, they had no opportunity to discover the new requirement until the formerly correct response had been weakened sufficiently by non-reinforcement. Lashley (10) attempted to create in the rat a general concept of reversal such that presentation of an external cue would immediately evoke a reversing response, but failed to generalize the concept to new situations. Had he succeeded, it would have been possible to test for the result reported here, since the animal's set could then be readily reversed. The rats would, however, have to be able, as assumed, to make the formerly correct response symbolically or partially so that it would provide the cue to the opposite response.

Hilgard and Humphreys (4) working with the conditioned eyeblink in human Ss found results which also appear out of line with those reported here. They conditioned Ss to respond positively to one light stimulus and not to respond at all to a negative stimulus, another light. This is comparable to the present situation in which the Ss were to respond to one word, and not to respond to another, successively presented. Then, as here, the Ss were asked to reverse their responses voluntarily by responding to the negative stimulus and inhibiting responses to the positive stimulus. They were more successful at producing responses to the negative stimulus than they were at inhibiting responses to the positive stimulus (4, p. 297). The explanation of the latter fact lies in the absence of one of the basic requirements for positive transfer to reversal training, namely, the ability to make the first-learned response symbolically, implicitly, or partially. A conditioned eyeblink can probably not be made symbolically because of its short latency and involuntary nature. Since the wink is the cue to making the reverse (inhibiting) response, it follows that the *S* could not successfully inhibit it before it had occurred, unless it *had* occurred symbolically. Therefore, the result obtained by Hilgard and Humphreys fits in with the present analysis.

Earlier it was noted how the present result apparently contradicted negative transfer theory. The explanation so far evolved will not only solve this difficulty, but it will also serve to clarify other situations which, though they follow the negative transfer paradigm, have yielded positive transfer. McGeoch concludes (13, pp. 427–28) on the basis of a number of studies that in situations designed to yield negative transfer, as the degree of training on the first task increases above that required on the test task, the negative transfer frequently gives way to positive transfer. This result is readily understandable according to the present analysis, since the

much stronger originally learned responses may be consistently inhibited as wrong, and this inhibition should serve to delimit the number of possible right responses in the new task and consequently to facilitate the choice of the correct responses. If, as in the present instance, there is only one alternative response, knowing what is wrong and being able not to do it will produce maximal positive transfer to a reversal task. As the number of alternative responses increases, the positive transfer should decrease.

A similar interpretation may also be made of negative practice data. The implication is clearly that negative practice may facilitate learning a right response by making a wrong response stronger, a possibility overlooked by Peak (**15**) in her otherwise thorough combing of possible explanations. The seeming paradox that strengthening the wrong response should aid learning a competing response is resolved if the first response is conceived as being used as an intermediate reaction providing a cue to the other response. That is, it is possible that practicing the spelling 'hte' strengthens it so that it is evoked the next time the person tried to spell the word, but since the new set marks this response as wrong, it will be inhibited and will serve as a cue to the correct spelling 'the.' The particular point of the present experiment is that it shows the *negative practice need not be done with the knowledge that it is wrong for positive transfer to result.* A reversal set or a set not to make the originally learned response will serve to make negative practice the means to positive transfer.

It should be noted that negative practice will be effective according to the present analysis only if the wrong response is initially as strong or stronger than the right response. If it were initially weaker, practice on it would only serve to make it more nearly equal in strength to the right response, neither response would appear consistently, and learning would suffer. The significant point is that practice, whether negative or positive, must increase the difference in response strengths of the two competing responses so that one or the other will appear consistently. If the wrong one in terms of the present instructions has been strengthened, it will serve, nevertheless, almost as well, since it may be used as a cue to its competitor. The general principle may be stated thus: *any practice which will increase the difference in strengths of two competing responses will be effective in facilitating the performance of either*, providing both responses are part of the organism's repertoire, both can be symbolically represented, and the organism can be given a set to reverse.[4]

[4] This statement may be readily deduced from the more general principle supported by all sorts of data that "in a conflict situation the probability of blocking becomes greater as the strengths of the competing responses approach equality" (**13**, p. 471).

Not all negative practice results need be explained by this principle. The knowledge that a mistake is being practiced may operate in any of the ways Peak (**15**) has suggested. But in some instances it would appear simpler to assume that the knowledge is irrelevant and that the improvement comes through the principle outlined here.

There is a further implication of the principle. If the wrong reaction is used, because it is stronger, as the cue to the right reaction, it should in time become short-circuited. That is, the right reactions, originally made through the wrong ones, should accumulate a certain amount of habit strength through direct reinforcement. This means that in time the right and wrong reactions will tend to become of nearly equal strength once more, a condition which should slow learning. To put it another way, as soon as the right responses are strengthened, they will tend to be evoked by the stimulus, but since the *S* is set to reverse, they will be reversed along with the wrong responses that still remain stronger. It follows from this that negative practice, though it may be temporarily beneficial, may later delay complete mastery.[5] There is some evidence in the present experiment that positive transfer was greater in the first part of learning the reverse discriminations than later but it is by no means conclusive. Siipola and Israel report results (**16**, p. 218) which definitely support this deduction, however. They found that with a high degree of original training positive transfer to the second task was initially very high but it decreased until the advantage of the initial training had disappeared by the fifth tenth of the learning of the second task, only to reappear again later.[6] Here again the conclusion is inescapable that the maximum interference between two tasks is greatest when they are equally learned. Strengthening either relative to the other will facilitate the performance of either.

Throughout this discussion reliance has been placed chiefly on the mechanism of the wrong response being used as an intermediate reaction as that seemed the most likely explanation of discrimination reversal results. Siipola and Israel (**16**) do not suggest a similar

[5] Hill and McLeod (**6**) found clear evidence of this effect in a retroactive inhibition experiment in which synonymous responses in the interpolated list were associated with the same stimuli as were used in the original list. When the interpolated list was learned to a much higher degree than the original, the interpolated intrusions early in recall of the original list were reliably fewer than they were later in recall. In other words, so long as the interpolated responses were stronger, as they were early in recall, they could be rejected as wrong consistently (and used as cues to the right responses, since they were synonymous with them). But as soon as the original responses reached a nearly equivalent strength, later in recall, the differentiation broke down and the interpolated responses tended to intrude. Strength was no longer a consistent cue to wrongness.

[6] More indirect evidence may be found elsewhere. For example, Melton's study of reminiscence in pursuit-learning (**14**), caused possibly by negative practice in the rest interval, showed such a decrement following the initial increment.

mechanism because they were apparently not interested in explaining the positive transfer they obtained. They state only that the stronger responses from the first task were often evoked without interfering with the performance of the second because it proceeded at such a different rate.[7] This suggests that there may be a second mechanism involved in positive transfer from negative practice, namely, that of decreased generalization. That is, increasing the difference in strength of two competing habits not only allows one in some cases to be used as a cue to the other, but it also serves to differentiate them so that they compete less. If they compete less, the practice which produced the differentiation, whether it was negative or positive, should facilitate the performance of either.

Summary

The experiment was designed to test at the human level the generalization based on animal studies that the difficulty of learning reversed discriminations increases with the amount of original learning.

Twenty-four *S*s learned to choose one of two responses to each of 15 stimulus words presented at the rate of one every two sec. They learned on four separate days, the first for practice, the other three for one of three work conditions arranged in a counterbalanced order. The conditions provided a control group which learned the same responses all the way through with 30-sec. rests after the criteria of 11 in 15 and 15 in 15 correct responses had been reached, and two reversal groups, one of which reversed after the criterion of 11/15 and the other after the criterion of 15/15 had been reached. The following generalizations were supported by the results:

1. As compared with an unreversed group, reversal after both criteria caused a slight but reliable decrement.

2. As compared with a group having no previous practice, reversal training up to both criteria produced large and reliable positive transfer.

3. In direct contrast with animal studies, the *ease* of learning reversed discriminations greatly increased with increasing amounts of original training.

4. This result was interpreted not as meaning that discrimination theories at the animal level are useless in dealing with human behavior, but rather as meaning that such theories will have to be extended to include such mechanisms as reversal sets and anticipatory responses to account for human behavior.

[7] Unlike the case in discrimination reversal, in their study rejecting the stronger wrong response would give the *S*s very little aid in making the right response. So the positive transfer obtained some other explanation.

5. The result was further interpreted as a case of negative practice, performed without the knowledge that it was wrong, which nevertheless yielded positive transfer. The explanation of the phenomenon was thought to lie in the use of the stronger wrong response as a cue to the right response.

6. These considerations led to the formulation of the general principle applicable at the human level that *any practice which will increase the difference in strengths of two competing responses will be effective in facilitating the performance of either.* In a psychological era which has insisted on the similarities between animal and human behavioral principles, it is worth emphasizing that here is one point at least where a principle derived at the animal level has to be radically revised to apply to humans.

(Manuscript received November 9, 1943)

Bibliography

1. Gulliksen, H. & Wolfle, D. L. A theory of learning and transfer. I. *Psychometrika*, 1938, **3**, 127–149. II. *Psychometrika*, 1938, **3**, 225–251.
2. Hilgard, E. R. An algebraic analysis of conditioned discrimination in man. *Psychol. Rev.*, 1938, **45**, 472–496.
3. Hilgard, E. R. & Campbell, A. A. Vincent curves of conditioning. *J. exp. Psychol.*, 1937, **21**, 310–319.
4. Hilgard, E. R. & Humphreys, L. G. The effect of supporting and antagonistic voluntary instructions on conditioned discrimination. *J. exp. Psychol.*, 1938, **22**, 291–304.
5. Hilgard, E. R. & Marquis, D. G. *Conditioning and learning.* New York: D. Appleton-Century, 1940.
6. Hill, C. J. & MacLeod, S. The effect of response similarity on retroactive inhibition. (Unpublished study)
7. Hull, C. L. Goal attraction and directing ideas conceived as habit phenomena. *Psychol. Rev.*, 1931, **38**, 487–506.
8. Krechevsky, I. A study of the continuity of the problem-solving process. *Psychol. Rev.*, 1938, **45**, 107–133.
9. Krechevsky, I. & Honzik, C. H. Fixation in the rat. *Univ. Calif. Publ. Psychol.*, 1932, **6**, 13–26.
10. Lashley, K. S. Conditional reactions in the rat. *J. Psychol.*, 1938, **6**, 311–324.
11. McClelland, D. C. Studies in serial verbal discrimination learning. III. The influence of difficulty on reminiscence in responses to right and wrong words. *J. exp. Psychol.*, 1943, **32**, 235–246.
12. McCulloch, T. L. & Pratt, J. G. A study of the pre-solution period in weight discrimination by white rats. *J. comp. Psychol.*, 1934, **18**, 271–290.
13. McGeoch, J. A. *The psychology of human learning.* New York: Longman's Green, 1942.
14. Melton, A. W. The effect of rest pauses on the acquisition of the pursuitmeter habit. *Psychol. Bull.*, 1941, **38**, 719.
15. Peak, H. Negative practice and theories of learning. *Psychol. Rev.*, 1941, **48**, 316–336.
16. Siipola, E. M. & Israel, H. E. Habit interference as dependent upon stage of training. *Amer. J. Psychol.*, 1933, **45**, 205–227.
17. Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43**, 427–449.
18. Spence, K. W. Continuous versus non-continuous interpretations of discrimination learning. *Psychol. Rev.*, 1940, **47**, 271–288.
19. Williams, S. B. Reversal learning after two degrees of training. *J. comp. Psychol.*, 1924, **34**, 353–360.
20. Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938.

THE GENETIC DEVELOPMENT OF PATTERNS OF VOLUNTARY ACTIVITY

BY R. C. DAVIS

Indiana University

In a series of reports the writer has presented evidence that psychological activities (those investigated, at least) are revealed in the organism by muscular activity, often of slight intensity, though of widespread distribution. Further, 'geographical' and temporal patterns of different sorts characterize the activity of the organism under different stimulating conditions. Much work from other sources has produced similar results (Courts **2**). The trend of the evidence points, it seems, to the general theory that psychological activities, so far as events internal to the organism are concerned, are characterized by complex patterns of muscular reaction, possibly vestigial remnants of overt movements. With such a general proposition in view it is clear that much interest lies in the genetic development of activity patterns for any particular sort of response. It is the purpose of the present experiment to study this point by comparing patterns shown by two different age groups in the same stimulus situation. For simplicity it was desirable to choose for study a situation which requires an actual movement from both adults and children: namely, the lifting of a given weight at a certain signal.

This sort of experiment therefore bears also on the topic of motor development, important in its own right. According to one view (Irwin, **8**) the 'original' form of motor response is 'mass activity' from which in the process of development, there is a 'differentiation' of more specific activities. Although Irwin (**9**) does not believe such a theory implies that later specific movements are 'contained in' a prior gross activity, there are nevertheless certain observations which indicate the existence of just such a sequence.[1] To generalize such a rule would certainly be questionable, but even if the specific merely succeeds the general response, it seems almost necessary that there be a stage where both are present, and probably a further development in which the general activity recedes. It is possible

[1] In Coghill's (1) description of locomotion in amblystoma it may be said that each segment actually participates in 'mass movement' before it moves independently. McGraw (12) reports the appearance of digital movements at the same stage as arm movements in her description of the reaching-prehensile movement.

that a movement develops by the addition of new elements as well as by subtraction of old. However this may be, in the earlier stages of development there is apparent involvement of a wider area of the organism in a response. But there is evidently only a relative loss of mass activity. Studies of action potentials already referred to show its presence in some degree in adults. With less mature individuals the patterns of muscular activity would be expected to show a less sharp focus. They might lie, one would suppose, somewhere between the adult condition and the easily detected 'mass activity' of infants.

In adults the development of action patterns with practice has been the subject of a number of studies, which have been summarized by Courts (**2**). The general trend of results is that tensions in remote parts (perhaps in active parts likewise) are diminished with practice, at least until a certain proficiency is reached. There may be a weakening of intensity, a sharpening of focus, or both.

Procedure

The performance chosen for study in this experiment was lifting in rotation a series of three different weights, one at the delivery of each visual signal. Each *S* was required to lie, ventral side down, on a cot in the darkened shielded booth. The head was always turned toward the right, and the arms were extended from the body with the elbows flexed 'wing fashion.' The left hand rested on the first of the weights to be lifted. Electrodes were connected and *S* was instructed to relax as completely as possible and to imagine that he was "going to sleep;" and that when a light came on he was to lift the first weight with a wrist motion and immediately set it down again and rest his hand on the next weight, thereupon repeating the performance until each weight had been lifted twice (four times for adults). The method of lifting was demonstrated.

Two independent amplifying and recording systems were used for measuring the muscle action potentials. The design and calibration of these instruments have been described in a previous paper (**5**). For adults, the single-sided type was used, while for the children the balanced circuits with ungrounded first stages were adopted. One, operating at a relatively low sensitivity, was always connected (active lead) to the dorsal crest of the left forearm and to a large thick pad on the left upperarm (neutral or reference lead). The other amplifier operating at higher gain was connected to various locations at different times and for different groups, as indicated below. For recording right arm activity a pair of electrodes similar to those on the left arm was used. When it was desired to record from either leg the active electrode was placed on the crest of the calf and the reference electrode on the ankle of the same leg. In recording from the jaw the active electrode was on the crest of the right masseter and the reference electrode around the neck. Inasmuch as it was desired to secure records from two of these 'remote' locations for some *S*s and from three for others, the electrode was changed to another location at the end of a series of lifts and the whole lifting performance was repeated for each new location of the electrodes.

Before any signal to lift was given a period of rest was allowed *S* while *E* observed the deflections on the face of the cathode ray tubes. From experience it was possible to estimate the approximate level of relaxation obtainable under the circumstances. When, as indicated by the oscilloscope, this was not being reached in the course of about five minutes, *E* went to *S* with reassurance, and further suggestions of sleep and relaxation. If such attempts did not succeed in the course of a few minutes the experiment was started anyway, for it was discovered that prolongation of the preliminaries produced restlessness instead of relaxation. The moving film camera was started just before the signal to lift was turned on, and stopped when the lift was completed. (This point could be detected from the deflection on the oscilloscopes.) *E* then waited for the return of approximately the initial level of rest (if possible) and thereupon delivered the next signal.

The adult *Ss* used in the experiment were 12 undergraduate students, results on whom were partially reported in an earlier publication (4). Weights used for these *Ss* were 100, 500 and 1000 gm. Records were taken in the manner stated from right arm, left leg, and right leg, while one recording system was always connected to the left arm. The order of recording and of lifting was reversed in half of the cases. The children used as *Ss* were 8, 9, and 10 years old, averaging about 9 years of age, taken from the third and fourth grades. Twenty-four useable records were first obtained from these *Ss*. It was thought advisable to lower the weight series somewhat to compensate for their lesser strength. Weights used, consequently were 50, 100 and 500 gm. Records were taken from the three remote locations in the same manner as they were for adults. This group of children will be designated Children I.

After their records were measured and averaged it was found (as will be discussed more fully later) that their level of 'rest' showed considerably more activity than that of the adults. This had in fact been suspected during the recording, but it seemed impossible to get them to do any better. Since it seemed possible that this difference in rest level had affected the critical results, a second series of readings was taken, from the same children as far as possible. The second sitting was about three months after the first. Since some were no longer available the group was filled up by using a few children who had been through the experiment but whose readings were not included in Children I. This second set of readings, designated Children II include 20 cases. In order to extend the field of information as well as check the previous readings the procedure was changed for this group. Electrodes were connected to the two arms as before, but in place of the two leg locations, a location on the crest of the right masseter was used. Another variation was a substitution of a 200 gm. weight for the 500 in the series to be lifted.

Records were measured in the manner described previously (4), representing the maximum potential during 1/6 sec. just prior to each lift and the maximum during the 1/6 sec. including the greatest excursion during lifting. The difference between these two measures was taken, and when converted into microvolts served as the fundamental representation of the amount of muscular action during the lift. For the children this same treatment was given to records from both 'active' and remote muscle groups. For the adults no subtraction was performed for the 'active' muscle group for the reason that no rest period activity was discernable at the sensitivity used. The next step in calculation was to combine the results for the lifts of the same weight for the 'active' regions and for the various remote regions.

Results

1. *Comparison of 'Rest Levels.'*—Since the measurements here designated 'rest level' were taken during the experiment it is clear that they should not be expected to represent any absolute rest level. It is probable that at the times these readings were taken the *Ss* were to some extent set for reacting to the signal and consequently were not entirely relaxed (3). For present purposes, however, such a set of figures will be seen to be the most desirable. A comparison of the averages for the rest level in the remote locations is given in Table I. The averages for adults are very similar to those reported for a larger group of *Ss* under similar conditions in a previous study. A striking difference is apparent between the level of rest for children and that for adults in spite of the special efforts made to secure relaxation in the children. Comparing first sittings, it is found that children showed about three times the amount of muscular activity found in adults. In their second sitting the children showed a significant decline in level, but are still very definitely more active than the adults. The bearing of these facts upon response characteristics will need to be considered.

At first glance such figures might be understood as meaning that children of this age, at least during their waking hours, are less relaxed than adults. Such may indeed be the case. Standard norms for basal metabolism (DuBois, **6**), it will be recalled, show a similar difference between adults and children. However there is the obvious difficulty of establishing an *S*'s condition under non-experimental circumstances when it is possible to measure him only

TABLE I

COMPARISON OF 'REST LEVEL' IN ADULTS AND CHILDREN (MICROVOLTS)

	Right Arm	Right Leg	Left Leg	Right Jaw
Adults: Ave.	4.35	4.02	4.00	
SD.	.92	.82	.93	
Children I: Ave.	14.55	12.32	11.11	
SD.	5.0	4.6	3.5	
Children II: Ave.	10.22			14.0
SD.	4.1			7.7
Children I – Adults				
Diff.	10.2	8.3	7.1	
t	6.7	5.8	6.8	
Children II – Adults				
Diff.	5.9			
t	4.8			
Children I – Children II				
Diff.	5.3*			
t	3.85			

* Considering only those used in both sittings (17 cases).

under laboratory conditions. If children are generally more stimulated by the experimental conditions than adults are, the difference in the rest levels would be accounted for. Being more familar with such situations, adults may be less disturbed. If, finally, it may be supposed that adults are likely to be better acquainted with the situations they encounter than children are, there would be a pervasive difference in so-called rest levels in the two groups, and the two interpretations would arrive at practically the same point.

2. *Responses of the Active Member* (*Left Arm*).—Readings on the active member for *S*s in the three experimental groups are summarized in Table II. Reading across the table there is evident a similar progression of activity in the three groups as the value of the weight lifted increases. These progressions are shown graphically in Fig. 1. There is in all three a suggestion of positive acceleration, though certainly little can be said of the shape of the curve on the bases of so few points, and they may well be linear, as certain theoretical considerations would imply.

In this connection an anomalous type of record may be significant. Six of the children tested (but none of the adults) showed a single large wave of potential, sometimes repeated, at the time of the

weight lift instead of the typical series of spikes of smaller magnitude and duration. Five of these *Ss* showed such large waves only at the beginning of the series, then gradually developed the typical response, while the other continued with the atypical form. Such a wave in the active member was paralleled by a similar though much smaller one in the remote parts. Since it was impossible to make comparable measurements on these records none of them were

TABLE II

ACTION POTENTIALS IN ACTIVE MEMBER (LEFT ARM) (MICROVOLTS)

Weight Lifted	50 gm.	100 gm.	200 gm.	500 gm.	1000 gm.
Adults: Ave.					
N = 12		124.2		149.6	191.2
SD		33.4		51.4	36.6
Children I: Ave.					
N = 24	177.2	179.8		244.0	
SD	78.2	94.6		77.4	
Children II: Ave.					
N = 20	166.5	173.5	216.5		
SD	101.5	79.0	106.5		
Diff., Children I – Adults		55.6		94.4	
t		2.0		3.8	
Diff., Children II – Adults		49.3			
t		2.0			
Diff., Children I – Children II	10.7	6.3			
t	.4*	.2*			

* Correlations between the two series are negligible.

included in the tabulation of results. Four of the six cases of anomalous waves occurred in first sittings.

The explanation of these waves is not at all certain. They may represent a synchronized discharge of all motor units correlated with undue haste and overexertion. A maximal load, therefore, might produce this kind of response.

The trends shown in Fig. 1 for children and adults, though substantially parallel, are at quite different levels. Reading down the columns of Table II one may compare the performances of the three groups at several weight values. For the 100 gm. lift which was common to all groups, the difference between adults and children is significant at the 1 percent level in both cases. Likewise, for the 500 gm. lift the difference between adults and Children I is highly significant. There is no doubt that children produce larger action potentials than adults in the active member while doing the same amount of physical work.

Explanation of this fact was sought in three directions. It was thought that the children might have performed the work with a quick jerk, producing a larger but briefer disturbance, whereas the

adults might by lifting gradually have generated just as much potential *in toto*, but spread it over a longer time. To test this hypothesis the records of adults and Children I were measured to discover the time between the beginning of the lift, as evidenced by the action potentials, and the point of maximum exertion. The

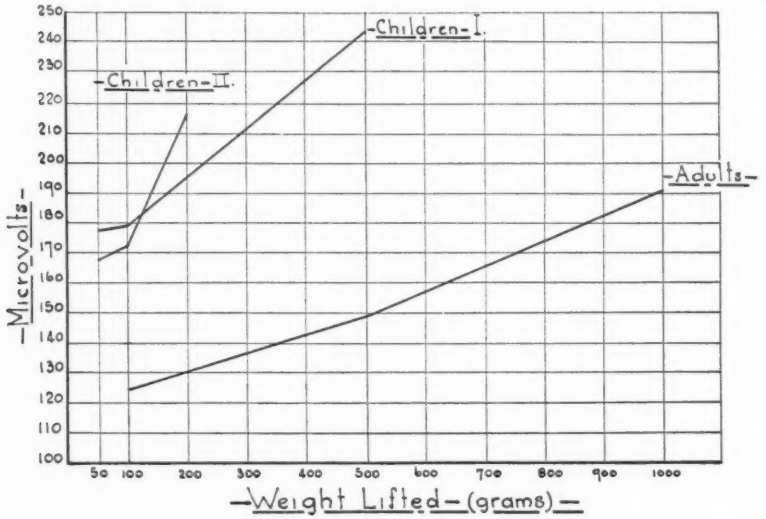


FIG. 1. Average response in active member for three experimental groups.

average times are shown in Table III. A systematic increase of time with weight may be noted, but the difference between adults and children in this respect is clearly non-significant, and speed of lifting as a differentiating factor must be dismissed.

In the second place it will be recalled that the experimental series for adults and children were not of the same length. Consequently

TABLE III

AVERAGE TIME FROM BEGINNING TO PEAK OF MUSCULAR RESPONSE (ACTIVE MEMBER)
Unit = 1/60 sec.

	Right Arm*				Right Leg				Left Leg			
Weight Lifted (gm.).....	50	100	500	1000	50	100	500	1000	50	100	500	1000
Adults.....		22.8	30	33.5		20.7	25.2	28.9		18.7	25.0	32.0
Children.....	19.3	22.0	25.1		19.6	21.6	25.7		17.3	19.3	24.7	

* The measurements here summarized were taken from the left arm. The headings merely designate the remote parts which happen to have been measured at the same time. The three series would be expected to show the same values and are useful simply as an indication of consistency.

a practice effect would influence one average more than the others. Investigation shows a small (about 10 percent) decrease in the adults' figures between the first and last series of lifts, and a similar increase for both groups of children. Both of these trends are non-significant in the present data. In any case they are certainly too small to account for much of the difference between adults and children.

The third possibility considered was that the explanation might lie in the difference in muscular size in the two sorts of S. The force exerted by such a muscle as here concerned is known to be proportional to its anatomical cross-sectional area (Gray's Anatomy, 10). This can be determined with reasonable accuracy from measurement of the circumference of the forearm. Such measurements were therefore made on Children II. None were obtainable on the particular adults used in the experiment, but as a substitute arm measurements were taken on a group of 72 adults of the same age range and sex distribution as those in the experiment, so that the average differs from that of the experimental group probably only by reason of sampling errors. Average circumference for this group was 9.47 in., and for Children II, 7.66 in. Cross-sectional areas would of course vary with respect to each other as the squares of these numbers, and the ratio between the squares is .73. We may compare this with the ratio (inverse) for action potentials, which is .67. Considering the approximations involved this seems excellent agreement, and the general rule may be stated that for equal amounts of physical work the size of the action potential varies inversely with the cross-sectional area of an individual's muscle.[2] Herein seems to lie a satisfactory explanation of the difference between adults and children in the responses of the active member. The total excitation must be the same for the same load; in the smaller muscle more of it must be carried by any given fiber group.

One further comparison appears in Table II. It may be seen that the response of the active member did not differ for the children from the first to the second sitting. (These figures it will be recalled represent activity over and above the rest level.) Since activity during 'rest' did decline, it may be inferred that the amount of activity during the lift is independent of rest level, within the present range, at least.

3. *Comparison of Remote Activities During Lifts.*—The principal question of the present study is of course, concerning the activity in remote parts. The immediate data on these are summarized in

[2] It might at first seem surprising that such a rule should obtain, since the difference in levers through which the force is applied seems to be neglected. These differences, however, would be of no effect if the ratio of the lever arms remains constant, as seems very nearly the case in this instance.

TABLE IV

ACTION POTENTIALS IN REMOTE MEMBERS (MICROVOLTS)

	Adults			Children I		
Weight (gm.)	100	500	1000	50	100	500
Right Arm: Av.	2.06	2.23	3.06	8.2	9.1	12.2
SD	2.65	1.95	2.65	5.5	8.5	8.8
%*	1.64	1.46	1.53	5.0	5.7	4.8
Left Leg: Av.	.59	.66	.73	4.0	4.2	8.0
SD	1.00	.75	.92	5.7	5.3	4.6
%*	.50	.46	.39	2.2	2.2	3.6
Right Leg: Av.	.76	.67	1.55	3.1	3.8	7.0
SD	1.13	1.62	1.51	5.0	4.4	5.3
%*	.57	.45	.63	1.7	2.1	2.7

Children II

Weight (gm.)	50	100	200
Right Arm: Av.	8.7	9.0	11.7
SD	7.6	7.4	11.7
%*	5.54	5.64	5.82
Jaw: Av.	12.2	11.0	13.4
SD	9.6	9.1	11.6
%*	6.94	5.86	5.68

* Percentage of the simultaneous action potential in the active member (LA).

DIFFERENCES

Children – Adults

Weight (gm.)	100	500
Right Arm I*		
Diff.	7.0	10.0
t	2.83	3.90
Left Leg I*		
Diff.	3.6	7.3
t	2.43	9.5
Right Leg I*		
Diff.	3.0	6.3
t	2.29	3.89
Right Arm II*		
Diff.	6.9	
t	3.0	

* Roman numeral designates group of children involved.

Table IV. In all of these figures, as previously mentioned, activity level prior to the lift has already been subtracted. The table shows, then, definite evidence of the existence of activity in all the remote parts tested in each of the groups.

Reading across the table, there is seen to be a regular increase of activity with increase in the size of the weight lifted, except in the case of one entry. This trend is graphed for the three groups for the right arm in Fig. 2. There seems to be a tendency toward a

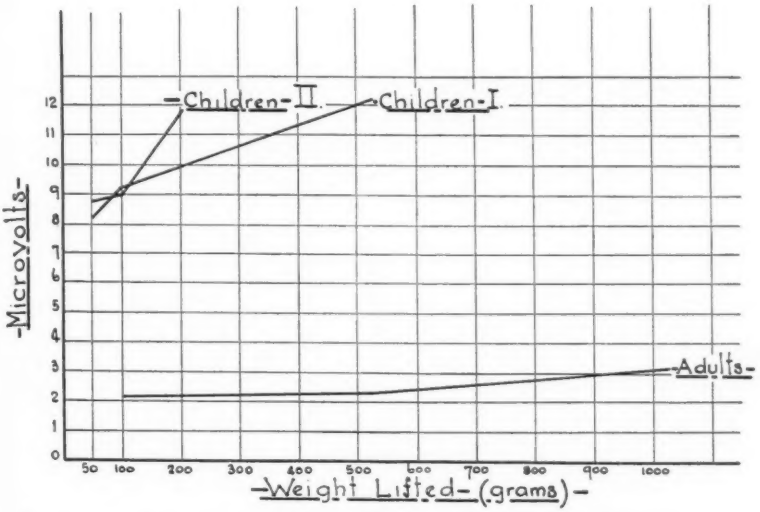


FIG. 2. Average response in a remote member (right arm) for three experimental groups.

somewhat steeper slope in the case of the children, as well as a higher level. (Practice effects in this function were studied by comparing similar lifts which occurred in the different positions in the experimental series. No consistent trends are to be seen within the limits of the present data.)

In making comparisons of the sizes of these remote activities with one another it is desirable to make some allowance for the magnitude of the simultaneous response in the active member. For this purpose the percentages shown in Table IV were computed. These were obtained by dividing the average remote response in each case by the corresponding average active member response, a technique which seems adequate for the purpose.

The pattern of limb excitation shows a distance gradient such as found previously, except that Children I does not show the excess of remote contralateral over remote ipsilateral previously found in adults.

The additional fact appears in Children II that activity in the right masseter is about the same as, or slightly larger than activity in the right extensor digitalis when the left arm is active. This can be filled into the gradient series very conveniently, though it may require a modified interpretation of the series. Patterns for the

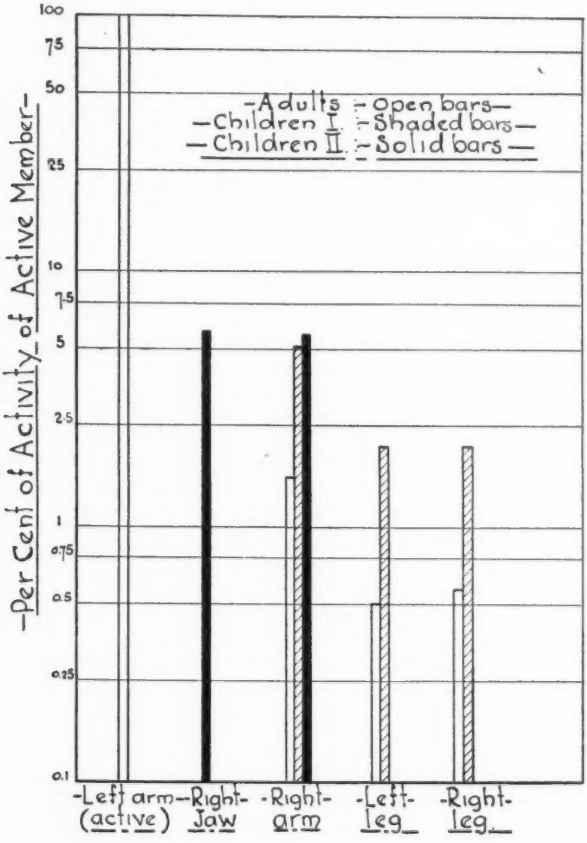


FIG. 3. The response gradient for the lifting of 100 gm. for three experimental groups.

several groups for the 100 gm. lift are shown in Fig. 3. (Note that ordinates of Fig. 3 are graphed logarithmically.)

A striking difference exists between remote excitations in children and adults, the averages for children being four or five times as large as those for adults. The differences in the averages (without allowance for active member activity) are evaluated statistically in

the last section of Table IV, in which it may be seen that they possess a high level of significance in general (though two are as low as the five percent level). The conclusion to be made from this table is that remote activity is more intense in nine-year-olds than in adults when the same amount of work is being performed.

So far the question remains, however, whether the children show any higher level of activity in remote parts than they would be expected to because of their greater active member excitation. The percentages computed above would suggest that they do, but such figures cannot well be evaluated statistically. It was intended to accomplish this evaluation by computing such a ratio for each individual for each situation and testing the significance of the average ratio. This procedure, however, proved unsatisfactory because of the low correlation existing between active member and remote excitations, particularly in one group. The effect was to exaggerate the differences between groups beyond reason.

The problem was therefore solved by use of analysis of co-variance. In brief that portion of the variance between children and adults which might be expected because of active member differences and their correlation with remote activities was subtracted, and the remainder of the intergroup variance evaluated in relation to the adjusted intra-group variance by means of the F-ratio. This was done seven times, as shown in Table V, once for each set of data to be compared.

Certain details of the procedure require comment. In the first place the experimental sub-groups (children and adults) are obviously unequal in their number of cases. The procedure here followed is essentially that recommended by Snedecor (**14**), which seems not unjustifiable mathematically in spite of the skeptical comment of Peters and Van Voorhis (**13**). It is true, of course, in this experiment that the number of active limb measurements is the same as the number of remote measurements for each sub-group.

The general procedure of the analysis is that described by Lindquist (**11**). That is to say, the adjusted inter-group variance was obtained by subtracting the adjusted intra-group from the adjusted total variance, rather than directly from group means.

Finally, the correlation used in computing the 'adjusted' values was the intra-group correlation rather than the correlation based on total populations. The reasoning behind this choice is given by Garrett and Zubin (**7**). (As a matter of fact calculations were made both ways. It was found that use of the 'total' correlation gave F-ratios that were somewhat lower but not enough different to change the conclusions.)

The F-ratios are all significant at the one percent level except one, which is significant at the five percent level and another which is somewhat less. The conclusion from this analysis is that when all due allowance has been made for the greater activity children show in the active member their remote excitations are significantly greater than those of adults. In other words, when the activity of the

TABLE V

COMPARISON OF REMOTE LIMB ACTIVITIES IN ADULTS AND CHILDREN WITH ALLOWANCE FOR ACTIVE LIMB ACTIVITIES (ANALYSIS OF COVARIANCE)

x-variable = active limb y-variable = remote location

Limb Position	RA	RA	LL	LL	RL	RL	RA*
Weight	100	500	100	500	100	500	100
Intra-group Σxy^2	180,621	198,674	235,588	457,972	417,872	293,768	210,341
Intra-group Σy^2	1,728	1,893	610	1,045	479	721	1,171
Intra-group Σxy	8,568	5,895	1,975	5,697	4,037	2,501	3,040
Intra-group r	.485	.306	.165	.260	.285	.173	.194
Adjusted Intra-group Σy^2	1,222	1,718	593	974	440	690	1,128
Total Σxy^2	207,261	283,754	281,400	504,725	436,700	391,607	325,388
Total Σy^2	2,115	2,692	712	1,471	538	1,033	1,531
Total Σxy	11,059	14,097	4,105	10,174	4,787	7,957	6,214
Adjusted total Σy^2	1,525	2,270	677	1,266	491	964	1,412
Adjusted Inter-group Σy^2	303	552	84	272	51	274	284
F-ratio	9.7	10.6	4.55	12.2	3.85	12.7	8.3
D.F.	1–32	1–33	1–32	1–32	1–33	1–32	1–29

* Figures from adults and from second sitting of children.

muscle of the left arm is the same in children and adults the remote excitations in children are greater.

We have in fact an experimental check on this statistical analysis, for the left arm activity of adults for the 1000 gm. lift is greater than the left arm activity of children for the 50 and 100 gm. lifts (see Fig. 1). Under these circumstances we find that the remote excitations in children for children lifting the lighter weights are two to four times as great as those in adults lifting the heavy weights.

Both methods of comparison show clearly that for comparable actions the remote excitations in nine year olds are much greater than in adults.

4. *Checks on Experimental Error.*—Inasmuch as the averages for children exceeded those for adults in all comparisons it was necessary

to consider with special care the possibility of some experimental error. The practical coincidence of figures for Children I and II for active member and remote activities and the expected drop in rest level on the second sitting give assurance that figures for children are at least self-consistent. During data collection the calibration of the amplifiers was frequently checked with a microvolter in the manner described elsewhere (**5**). Upon completion of the series calibration of the microvolter was itself rechecked by reference to a General Radio microvolter and found to be unchanged.

It will be recalled that the readings on adults were taken with a somewhat different amplifier circuit. This had been calibrated repeatedly at the time, but to check on the possibility of some undiscovered difference between circuits a short series of measurements was taken on a new group of 10 adults using the newer form of apparatus, which had been employed for the children. Averages for rest level, active member response, and for remote excitation in one location were obtained. They differed slightly from the averages of the actual experimental adult group (two being greater, one less) but no more than would be expected by reason of sampling error, as indicated by the *t*-test. The original figures therefore seem to be valid.

Discussion

In the three functions measured, rest level, active member response, and remote excitation there are differences that are referable primarily at least to age differences. Nevertheless it is possible that the three relationships exist for different reasons. In connection with the active member differences the evidence presented gave good reason to believe that the underlying factor is simply that of muscle size, which would not necessarily imply anything about habitual activity level. We may inquire then whether the same factor might explain the difference in rest level. But the difference between the rest levels of the two groups is far greater than would be accounted for on the basis of the size ratio, and, in the second place, to allege such an explanation would be merely to say that children compensate for their smaller muscular size by maintaining greater muscular tension. (As a matter of fact, speaking in these terms, they more than compensate.) It will be recalled, furthermore, that the rest level measures in children were affected by habituation to the situation, whereas the other two functions were not. It seems then, that in the higher rest level of the children we have a functional characteristic of the age group, either lower threshold of stimulation or higher level of autonomous activity. This changes with age, in part, no doubt, by reason of habituation.

In the case of the remote excitation differences it is quite likely that muscle size (active member) does play a role. However, substantial differences remain when this factor, at least in so far as represented by action potentials, is eliminated. The excitation pattern for comparable actions was seen to be, not necessarily more extensive in children, but more intensive, having less of a focus and a shallower gradient away from the focus. If 'mass action' in infants means action with little if any focus we see in the *S*s of this experiment two stages in the development of differentiation from this sort of response. Along with this sort of development there may also take place a shift in the position of the focus, particularly during the earlier years. There is, in fact, the possibility that the focus has shifted between the two age levels which were used herein, since a focus can be established only with reference to the points actually measured. Such a change in this performance in the interval concerned seems, however, rather unlikely, since the same muscle seems to be doing most of the work in both cases.

In a sense the excess of remote excitations in children might be referred to as another instance of their presumed greater excitability. This is probably not quite the same sense as that used in explaining rest level differences. In the one case the meaning intended is 'more frequent and larger responses,' in the other 'more massive responses.'

The reasons for the increased differentiation of response are, of course, not revealed in the present data. As mentioned earlier there is evidence that practice sharpens focus, though the limited amount here used did not. There is also the possibility of some growth process (maturation) independent of specific or even general practice. It might in addition be suggested that the change is symptomatic of a general modification of personality toward a greater detachment from the immediate situation.

To revert to the broad topic of the genetics of psychological processes introduced at the beginning of the paper, we may be led to certain inferences from the present results. If greater massivity is characteristic of all psychological processes in younger individuals there is room for less clear-cut distinction between one process and another. A movement of the right hand, to take a simple instance, would be more like a movement of the left in a child than in an adult, and in general one psychological process in the child would be more like another than would be the case with an adult. This expectation seems to be confirmed by general observations on children's overt activities and inferences about their hidden processes.

Summary

The genetic development of muscular patterns of voluntary response was studied by comparing excitations (action potentials) at various locations during weight lifting by adults and by children averaging nine years of age. The children were measured a second time after an interval of several months.

'Rest level' activity was found to be considerably higher in the children than in adults. At the second sitting it had decreased considerably but was still more than twice as great as that found in adults. The facts are taken to signify a regularly higher activity level in children.

The excitation of the active member increases as a function of weight lifted, but the level of activity is always much greater for children than for adults performing the same work. This difference is shown to be inversely proportional to muscle size. This segment of the response in children is unaffected by adaptation so far as the experiment goes.

The pattern of remote excitations in other limbs follows the same rules of focus and distance gradient as previously described for adults. In addition the excitation of the masseter muscle (contralateral) is shown to be equal to or somewhat greater than that of the contralateral forearm muscle. Remote excitations in the children are, however, much greater for the same weight lift than for adults. This is also true, as shown by a series of co-variance analyses and by experimental comparison when the remote excitations are compared on the basis of the same active member activity. The results may be described as showing a distance gradient whose slope is much less steep in the case of children.

These facts are interpreted as evidence of a declining massivity of response with increasing age, and as implying a less sharp differentiation between responses in children than in adults.

(Manuscript received October 30, 1943)

References

1. Coghill, G. E. *Anatomy and the problem of behavior.* Cambridge: Cambridge Univ. Press, 1929. Pp. xii + 113.
2. Courts, F. A. Muscular tension and performance. *Psychol. Bull.*, 1942, **39**, 347–367.
3. Davis, R. C. Set and muscular tension. *Ind. Univ. Publ. Sci. Sr.*, 1940, No. 10.
4. Davis, R. C. The pattern of muscular action in simple voluntary movement. *J. exp. Psychol.*, 1942, **31**, 347–366.
5. Davis, R. C. The design and testing of multiple amplifiers for action potential recording. *J. exp. Psychol.*, 1943, **32**, 270–274.
6. DuBois, E. F. *Basal metabolism in health and disease.* (3d ed.) Philadelphia: Lea & Febiger, 1936. Pp. 494.
7. Garrett, H. E., & Zubin, J. The analysis of variance in psychological research. *Psychol. Bull.*, 1943, **40**, 233–267.

8. IRWIN, O. C. The organismic hypothesis. *Psychol. Rev.*, 1932, **39**, 128–146; 139–202; 387–393.
9. IRWIN, O. C. Dennis on mass activity: a reply. *Psychol. Rev.*, 1933, **40**, 215–219.
10. LEWIS, W. *Gray's anatomy, descriptive and applied.* (21st ed.) Philadelphia: Lea & Febiger, 1924. Pp. 1249.
11. LINDQUIST, E. F. *Statistical analysis in educational research.* New York & Boston: Houghton Mifflin, 1940. Pp. xi + 266.
12. MCGRAW, M. B. Neural maturation as exemplified in the reaching-prehensile behavior of the human infant. *J. Psychol.*, 1941, **11**, 127–141.
13. PETERS, C. C., & VAN VOORHIS, W. R. *Statistical procedures and their mathematical bases.* New York: McGraw-Hill, 1940. Pp. 516.
14. SNEDECOR, G. W. *Statistical methods.* Ames, Iowa: Collegiate Press Inc., 1937. Pp. xiii + 341.

EXPERIMENTAL STUDIES OF THE JUDGMENTAL THEORY OF FEELING: VI. CONCRETE VERSUS ABSTRACT SETS IN THE PREFERENCE JUDGMENTS OF PICTURES

BY HENRY N. PETERS

University of Missouri

In the preceding paper of this series it was reported that the enhancing effect of learned positive reactions on the affective values of single colors occurred in *S*s who observed with a perceptual, or concrete, set, and did not occur in *S*s who observed with a conceptual, or abstract, set (**16**). The experiment to be reported in this paper was performed before the one described in the preceding paper, and was thus planned and carried out prior to the discovery of the vital role played by the *S*'s observing set. The results of this experiment had been examined thoroughly, lumping all *S*s together, as in the experiments reported in papers I to IV, and the conclusion that learned reactions did not measurably alter the values of pictures seemed evident. However, the data of this experiment were re-examined, segregating the Test *S*s according to their sets, and evidence was found for the conclusion that learned positive reactions do alter the values of pictures according to the prediction of the Judgmental Theory.

Procedure

Test *S*s came individually for two experimental sittings which were separated by a week's interval. During the first sitting they solved a learning problem which required them to make positive responses to five out of 20 picture postcards. In the second period they rated the pleasantness of 50 postcards, the 20 used in the learning problem and 30 additional ones.

The learning problem, which was defined for *S* in typewritten instructions, required the *S* to discover which one of four picture postcards to pull toward him whenever a lot of four was presented. A trial consisted of the presentation of five lots of four pictures. The lots were presented in the same order each trial, but the spatial arrangement of the four pictures of a lot was shuffled from trial to trial. *S* and *E* sat on opposite sides of a table, facing a gray screen, three feet high, which rested on the table. There was a narrow opening, with shutter, at the bottom of the screen directly in front of *S*. It was through this opening that *E* pushed a 12 × 17 in. board along the top of which were arranged the four pictures of a lot. The board was raised slightly at the top, the side away from *S*, so that it made an angle of about three degrees with the table top. It was covered with dark gray pasteboard. *S* was allowed one min. in which to choose a postcard and to pull it toward him to the bottom of the board. He was permitted only one choice per lot and if he had not made his choice at the end of 45 sec., *E* gave a warning signal. However, regardless of how quickly *S* made his choice, the board was not pulled back by *E* to his

side of the screen until the minute was up. When *E* had removed a lot, he checked *S*'s choice and said 'Right' or 'Wrong.' The problem was considered solved when *S* had made three successive errorless trials. This criterion was explained to *S* in the beginning.

E observed *S*'s movements through a mirror located high on the wall in back of *S*. He kept a record of trials, errors, and the particular pictures chosen at each trial. *E* had two of the boards which were used in presenting the pictures. By arranging the pictures of one lot while *S* was making his choice from the preceding one, *E* was able to present the lots in immediate succession. *S* was not allowed to handle the postcards in any way, except when he made his choice, and then he was required to touch only the narrow strip of Scotch drafting tape which was stuck along the bottom (always a narrow) edge of the card. This tape served also to cover the only printing on the surface of the cards.

Certain criteria were used in selecting the postcards. Every picture was printed in color. Many were portraits; all of them were predominately pictures of human beings, usually only one, at most 3. Pictures of obvious religious interest were eliminated. All were printed length-wise on the card. There was no writing or print visible on any of them. None was done in a 'modern' style; all were by recognized masters. With the exception of one American, all were European. Every print was of a picture conventionally recognized as 'good' art.

From the 50 pictures used in the entire experiment, the 20 used in the learning problem were selected at random. The five 'positive pictures,' those *S* was required to choose in learning, were likewise selected at random.

At the second experimental sitting, *S* was seated at a table on which rested a $6 \times 9 \times 20$ in. sorting box and the stack of 50 pictures. He was given the following typewritten instructions.

"INSTRUCTIONS FOR RATING THE PLEASANTNESS OF PICTURES

"On the table before you is a stack of postcard reproductions of oil paintings. Will you take these cards one at a time, in the order in which they are stacked, and examine each carefully, making a judgment of its pleasantness? You will make these judgments in terms of 7 grades:

(1) Very pleasant
(2) Pleasant
(3) Slightly pleasant
(4) Medium, that is, indifferent
(5) Slightly unpleasant
(6) Unpleasant
(7) Very unpleasant

"You will notice that the box on the table has 7 slots in the top, one for each of the 7 types of judgments. As you make one of the judgments for a card drop the card in the appropriate slot.

"Please make these judgments on the basis of your own private opinion of the pictures. In forming your opinions it may help to ask yourself some question such as the following one: How would I like to have a large copy of this picture to hang on the wall of my room? Do not use as your basis of rating the pictures what you suspect art critics might think of them. In this experiment we are interested in your own preferences, not in artistic values.

"Please handle the cards with care. Do not bend or mark them in any way. It will probably be best if you handle each picture by the bottom edge where a piece of tape has been glued.

"You may take as long or as short a time as you wish."

The stack of pictures was thoroughly shuffled before *S* entered the room. The slot at the right end of the box was the P extreme. Below each slot was pasted a piece of paper with the name of the rating, 'Very Pleasant,' 'Pleasant,' etc., typed on it. When *S* dropped a card in a slot it disappeared from sight. Thus *S* was not aware of the number of ratings of the various categories he was making.

When *S* had finished sorting he was given a sheet of paper on which he was requested to write an account (1) of the basis of his judgments of the pictures and (2) of any training he had had in art.

*S*s of the Control Group came only once and rated the 50 pictures. The conditions of rating were the same as those used with the Test Group. A verbal report was not taken with *S*s of the Control Group.

Forty-one men and 21 women were used as *Ss* in both Test and Control Groups. They were all typical freshmen and sophomores taken from introductory psychology classes. All were naive to psychological experimentation. None was aware of the nature of the problem. None of them had had any serious training in art.

Results

The majority of Test *Ss* found the learning problem an easy one. Mean trial and error scores for the group of 62 *Ss* are 4.79 ± 0.31 (σ_m) and 11.58 ± 0.77. The three criterial trials were not included in the individual scores from which the trial mean was derived. A count of the number of times each picture in a lot was reacted to (pulled) by each *S* discloses that in the great majority of cases the 'correct' picture received more reactions than any other picture in the lot. There were only five exceptions, out of a possible 310 (5×62), in which another picture received more responses. However, the important point is that every *S* was required in the end to make consistently what is defined herein as a positive response, i.e., touching and pulling, to only the five positive pictures.

This experiment, like the others described in this series, was planned to test a hypothesis which has been referred to as the Judgmental Theory. According to this theory when *S* makes a value judgment of an object he is signalizing the existence within himself of a positive or negative reaction tendency to the object. When his judgment is P, his reaction tendency is positive; when his judgment is U, it is negative. From this it has been inferred that if experimentally determined reaction tendencies, defined as positive and negative to specific objects, are established in *S*, they should be reflected in his affective judgments of those objects. This deduction has been verified for Japanese words and for single colors (**15**, **16**). In the case of colors, however, the expected enhancing effect of positive reactions (and conversely the depressing effect of negative reactions) was found to be conditioned by *S*'s observing set. The effect was evident only in the judgments of the group of *Ss* who observed under a relatively concrete set. This is what was found in the present experiment with picture postcards.

The first four columns of figures in Table I present the median ratings, and their Qs, of the five positive pictures for the whole Control Group and the whole Test Group. In tabulating the ratings the extreme U rating was listed as 1, the extreme P as 7. Although four out of the five Test medians are slightly higher than the corresponding medians of the Control Group, and thus show the trend expected on the basis of the hypothesis, the difference between the two groups is neither impressive to the eye nor statistically reliable. The median rating for all five positive pictures together is 4.1 for the Control, and 4.3 for the Test Group. The Qs for both of these

medians are the same, 1.2. The percentages of Favorable judgments are likewise similar for the two groups, for the five positive pictures. This percentage was found by dividing the sum of ratings in categories 5, 6, and 7 plus 1/2 the ratings in category 4 by the total number of ratings, which was 310 for both groups. This percentage was 52.6 for the Control, 56.9 for the Test Group.

TABLE I

MEDIAN RATINGS OF THE FIVE POSITIVE PICTURES BY VARIOUS GROUPS

Picture	Control Gr.		Test Gr.		Concrete Set		Abstract Set		Doubtful Set	
	Mdn	Q	Mdn	Q	Mdn	Q	Mdn	Q	Mdn	Q
I	5.1	1.3	5.2	1.3	5.0	1.2	5.8	1.0	4.5	1.4
II	4.1	1.2	3.9	1.2	4.4	1.2	3.6	1.2	3.7	1.2
III	4.0	1.2	4.5	1.3	5.3	1.1	4.1	1.6	4.4	1.2
IV	4.0	1.1	4.3	1.0	4.9	0.9	4.1	1.2	4.1	1.0
V	3.6	1.0	3.9	0.9	4.0	0.4	3.7	1.1	3.9	0.9

There was even less evidence of a true difference in the ratings of the two groups for the 15 'negative pictures,' the pictures used in learning but which *S* was not to react to positively. This will be evident to the reader when the data presented in Table II are described.

On the basis of their written verbal reports the writer sorted the *S*s of the Test Group into three sub-groups, according to whether, in his opinion, they had observed the pictures under (1) a concrete set, (2) an abstract set, or (3) an indecisive, or doubtful, set. In the case of each *S* attempt was made to answer this question: Were this *S*'s judgmental reactions to the visual pictures themselves (concrete set) or to meaningful concepts aroused by the pictures (abstract set)?

The following is an example of a report which was considered to portray an abstract set.

"I based my judgment on the personal appeal the picture had for me. Not necessarily on the painting but on the subject itself. If I liked the type of subject and its appearance I liked the picture. Also if the picture suggested some particular phase of life that appealed to me I thought it was pleasant in varying degrees."

Two examples of concrete set reports are:

"I don't really know. I just selected those which sort of struck my eye as being a little more colorsome than the others, and those which had more life to them."

"First, proportion and symmetry of the pictures. Second, the degree in which they impressed me as having the quality of warmth

and life. I prefer pictures which are neither too stiff and unemotional, nor too florid and highly emotional."

Reports which either gave reasons about equally balanced in frequency between the concrete and the abstract or were so inarticulate that it was impossible to decide on the abstractness of the set

TABLE II

PERCENTAGES OF THE RATINGS WHICH FELL IN VARIOUS CATEGORIES FOR THE CONTROL GROUP AND THREE SUB-GROUPS OF THE TEST GROUP

(1) 5 Positive Pictures

Group	N	Ratings					Favorable
		1 or 2	3	4	5	6 or 7	
Control Group	310	14.5	20.0	25.8	15.8	23.9	52.6
Concrete Set	85	7.1	9.4	32.9	16.5	34.1	67.1
Abstract Set	95	13.7	22.1	21.1	13.7	29.5	53.7
Doubtful Set	130	13.1	20.8	26.9	16.2	23.1	52.7

(2) 15 Negative Pictures

Group	N	Ratings					Favorable
		1 or 2	3	4	5	6 or 7	
Control Group	930	20.3	14.9	24.0	16.5	24.3	52.7
Concrete Set	255	13.7	14.9	28.6	20.0	22.7	57.1
Abstract Set	285	14.4	21.1	26.3	17.5	20.7	51.4
Doubtful Set	390	16.7	14.9	27.9	17.4	23.1	54.5

(3) 30 Neutral Pictures

Group	N	Ratings					Favorable
		1 or 2	3	4	5	6 or 7	
Control Group	1860	16.0	14.6	26.3	17.7	25.4	56.2
Concrete Set	510	13.3	15.7	30.4	15.5	25.1	55.8
Abstract Set	570	14.7	16.5	24.7	20.5	23.5	56.4
Doubtful Set	780	17.8	12.8	31.3	18.6	19.5	53.7

were thrown into the 'Doubtful Group.' In general the abstract set *Ss* were more articulate than the concrete. In this, as in previous studies in which verbal reports were taken, no *S* reported an influence of the learning problem on his judgments.

Seventeen *Ss* were judged to be concrete in their set, 19 abstract, and 26 doubtful. The numbers of men in these groups were 12, 14, and 15, respectively.

The last six columns of figures in Table I contain the medians and Qs of the ratings of the three sub-groups. With the exception of Picture I, the medians of the Concrete Set *Ss* are higher than the

corresponding medians of both Abstract and Doubtful Set *S*s and are, with the same exception, noticeably higher than those of the Control Group. On the other hand, the medians of the Abstract and Doubtful groups, while in four cases each exceeding those of the Control Group, are certainly more closely in agreement with the Control medians than are those of the Concrete Set group.

The differences between medians in this table were found to be significantly reliable in only one case, the difference between the medians of Concrete Set group and the Control Groups for Picture III. $D/P.E._D$ was found to be 3.35, for which P = 0.01.

Medians of the Control Group and the three sub-groups of the Test Group were also calculated for the other 45 pictures and the direction and magnitude of the differences between medians of the Control Group and the three sub-groups examined. While of the five positive pictures, four of the Concrete medians exceeded the Control medians (as shown in Table I), this was the case for only nine of the 15 negative pictures and 14 of the 30 'neutral pictures,' ones not used in learning. While four of the five Concrete medians for the positive pictures exceeded the corresponding median of the Control Group by 0.3 or more, this was the case with only four of the 15 negative pictures and five of the 30 neutral pictures. Two of the Concrete positive picture medians exceeded the corresponding Control medians by 0.9, while this is not true of any of the other 45 pairs of medians.

The second through the sixth columns of figures in the first part of Table II show percentages of the total number of ratings given the five positive pictures by the Control Group and the three sub-groups of the Test Group. Each percentage tells the portion of ratings falling in one of five categories. The categories are simply the seven divisions of the rating scale, grouping together the two divisions at each extreme. The other two parts of the table show the same percentages for the 15 negative pictures, and the 30 neutral pictures. The figures in column N show the total numbers of ratings by each group of *S*s. Thus, since there are five positive pictures and 62 *S*s in the Control Group, N = 310 for the Control Group in the first part.

It is apparent from visual inspection of this table that there is only one row of percentages, those of the Concrete Set *S*s in part (1), which differ in a consistent manner from those of the Control Group, and in a manner predicted by the Judgmental Theory. The percentages of the two U categories, 1 or 2 and 3, for the Concrete Set group are both lower than the corresponding percentages of the Control Group and the percentages of the two P categories are higher for the Concrete Set group. This trend is not shown by the

differences between Concrete and Control percentages in either of the other two parts, and it is not shown by the differences between Abstract (or Doubtful) percentages and Control in any of the three parts of this table.

The reliability of this trend in differences between the frequencies of ratings of the five positive pictures (1) by the Control and Concrete *Ss*, and (2) by the Abstract plus Doubtful *Ss* and the Concrete group, was tested in two ways. One was a chi-square test of the difference in distributions of the ratings over the five categories shown in Table II. The distributions for the Concrete and Control Groups yielded a chi-square of 9.39, which for $n = 4$ has a P of almost exactly 0.05. Grouping the Abstract and Doubtful *Ss*' ratings together and comparing the distribution with that of the Concrete *Ss* gave a chi-square of 8.44, which has a P between 0.10 and 0.05. In calculating these chi-squares Yates' correction for continuity was made (**10**).

The second method of testing the reliability of the difference between the distributions of these ratings yielded more convincing evidence of a true difference. The ratings of each group were divided in the two categories, Favorable and Unfavorable. The method of determining total Favorable ratings has already been described. The last column of Table II shows the percentage of Favorable ratings for every group. For purposes of calculating chi-square for the difference between any two groups a fourfold table was thus employed. Chi-square for the difference between Concrete Set and Control Groups is 5.67 which, for $n = 1$, has a P between 0.02 and 0.01. Chi-square for the difference between Abstract and Doubtful *Ss*, lumped, and the Concrete Set group is 4.90, with a P between 0.05 and 0.02.

There are two explanations, both of which are probably valid, of the failure of the data of this experiment to show a depression in the affective values of the 15 negative pictures. It was found in a previous experiment (**15**) that after a week the effect of negative reactions on affective judgments had almost vanished. Perhaps in this experiment the depressing effect had been present, but had entirely disappeared by the end of a week. The other explanation lies in the fact that in this study negative responses were not attached to single, particular objects, as they had been in the previous experiments. Here *S*'s attention was centered on the picture-to-be-responded-to by pulling, and any negative or avoidance response which might have been present was necessarily spread thinly over the three remaining pictures in the lot. This condition of the present experiment is one which would be expected to dilute any negative response effect.

Discussion

The results of this and the preceding experiment indicate that the shift in affective value which follows the learning of positive and negative responses is a function of *S*'s observing set. Specifically, when the set is one which points *S*'s attention to perceptual (concrete, sensory, patterned) features of the stimulus, the shift in value after learning is reliably measurable; when *S*'s set causes him to attend to conceptual (abstract, ideational, associative) derivations from the stimulus, the shift in value is not found. This functional relationship has been demonstrated to hold for single colors and for pictures; it has not yet been demonstrated for Japanese words, the materials used in the other experiments reported in this series.

The fact that there are many different kinds of observing sets has always been remarked by students of affective judgments. Although many have defined more than just two sets, the different varieties can be readily classified as primarily of a lower (sensory) level of abstraction or primarily of a higher (ideational) level of abstraction. This common difference between observing sets, i.e., their concreteness versus their abstractness, has been concealed by a rich diversity of labels which the different schools of investigators have given them.

Bullough (**2**) was first in the general field of affection to give perceptual sets careful study and definition. He considered the observing sets to be four in number, and called them the *objective*, *physiological*, *associative*, and *character* sets. The first of these is described in terms which imply that attention is dominated by stimulus qualities; the second set selects organic sensations for the focus of attention; while with the third set the judgment is really a reaction to ideas aroused by the stimulus object. The first two can readily be conceived of as two sub-types of the genus, concrete set; whereas the third is the most common variety of abstract set. The status of the 'character' set with respect to abstractness is more dubious. However, the writer is of the opinion that it is a relatively rare type of abstract set. Nakashima (**13**) identified two types of judgments in his *S*s, 'direct' and 'associative,' and explained them in terms of two attitudes, which he called the 'passive' and the 'critical' attitudes. His published protocols contain an excellent detailed description of the nature of the two types of judgments. According to the passive set, or attitude, "the affective judgment is prompted directly by the external stimulus"; whereas when the critical set is dominant, an inferential basis for judgment is sought. He reported that with increasing practice at making affective judgments, the critical set became less, the passive more, frequent.

Feasy (**5**) labelled the concrete and abstract sets, operating in the judgment of geometrical forms, the 'pattern' and 'picture' types. Yokoyama (**23**) described the two sets as the 'sensorial attitude' and the 'objective attitude.' Wells (**20**) identified three observing sets: the 'critical affective' and the 'critical perceptive,' which can be considered the same as Bullough's physiological and objective types, and the 'common sense attitude,' which compared with the first two is an abstract set. In her experiment these sets were established and controlled by instructions. However, Aveling's results (**1**) show that even with highly trained observers the abstractness of set is not entirely subject to instructional control.

It has been clearly demonstrated that the abstractness of observing set is a condition of the relative values of a group of stimuli. Pratt (**17**), for example, found that five *S*s' judgments of the relative pleasantness of tonal intervals occurred in two distinct orders. Investigation revealed that two *S*s were basing their judgments on *smoothness* of the intervals (concrete set), while three *S*s were using *musical meaning* (abstract set) as basis of judgment. Two investigators (**19, 11**) have reported a difference in readiness of judgments of the two types of sets to shift with experimental conditions. In both instances this difference is one which, like the difference found in the present study, can be interpreted as evidence of the rigidity of affective values governed by an abstract set as opposed to the lability of values governed by a concrete set. For three of Valentine's five *S*s repeated stimulation resulted in increased pleasantness of discords. Two *S*s who failed to show this shift were found to be observing under a 'critical attitude,' or discriminating the musical meanings of the intervals. In a study of affective contrast between modalities, Harris found that very P (or U) odors interspersed with colors resulted in a tendency for affective values of the colors to shift toward the affective pole of the odors. In other words, contiguity and not contrast holds between modalities. However, his *S*s' introspective reports revealed that the shift in affective value of the colors did not occur when they were observing under a 'critical attitude.'

A tendency for *S*s' preference judgments to fall into two types has been reported in a series of English studies (**3, 4, 21**). The terms which are employed by these investigators to characterize the two types strongly suggest the concrete versus the abstract set. In these studies a group of *S*s' judgments of a class of objects were factor analyzed. In addition to a general factor, a bipolar factor was found for every class of objects used. This bipolar factor, with which some *S*s are positively saturated, others negatively saturated, is the evidence presented for the types. The nature of the types

was determined by examination of stimulus objects on which the preferences of the two types differed widely, and by comparison of introspective reports of individuals of the two types. For picture postcards, Dewar characterized one pole as preference for the 'objective-classical,' the other pole as preference for the 'subjective-romantic.' Williams, Winter, and Woods explained this type tendency in judgments of literary compositions as a preference for 'classical style' versus preference for the 'romantic school.' They quote the following words of Burt in support of their description of the nature of the two types. "We might perhaps summarize each type as follows. The comments (of the *S*s) show that their chief interest lies: (1) in the thing actually pictured or described (the objective type); (2) not in the thing itself, but in its effects upon themselves (the subjective type). . . ." Eysenck interpreted the bipolar factor in picture preferences in terms of a tendency to favor 'formal' and a tendency to favor 'representative' types of art. It is very probable that underlying these differences in preferences is variation in abstractness of sets.

The tendency for different *S*s to react on different levels of abstraction has attracted the attention of students of perception itself, as well as investigators who were primarily concerned with affective attitudes. Granit (**9**) used the method of brief exposure with figures of varying degrees of complexity and familiarity. His *S*s were required to reproduce the figures. He remarked that *S*s tended clearly to fall into two types, an 'objective' and a 'subjective.' The first showed cautious dependence upon the stimulus, seldom resorting to construction (a concrete set); whereas the second type readily transformed the primary impressions, treated the elements freely, and frequently resorted to past experience (an abstract set). Perhaps the most convincing account of the fundamental nature of these two modes of observing is given by Goldstein (**8**). He was first struck with the extreme concrete type of perceiving in working with brain pathology cases. He also found that it is possible for some normal *S*s to maintain either a concrete or an abstract set in observing. Goldstein stressed the importance of carefully identifying the nature of *S*'s perceptual set as a prerequisite to adequate interpretation of his perceptual reactions. It may be added that, since all reactions with the exception of purely reflex ones are perceptual reactions, establishment of *S*'s level of abstraction is vital to interpretation of any part of his adaptive behavior.

Five modes of perceiving are described by Neff (**14**) in terms which clearly reveal that the essential difference between the modes is one of degree of abstraction from the sensuous object. "The *figurational* . . . apprehension of a figure as an incorporeal line

drawing . . . and *depictive* . . . apprehension of a figure as a concrete, particular object . . . forms closely adhere to the region of the object. The last three, however, depart from that sphere and strike out in increasing measure for new psychological regions. The *abstractive* variety takes the object as a representative of some other object; the *generalized* mode carries through a topical kind of reference, the object standing for a class, and the *symbolical* mode regards the object merely as a surrogate, as a 'sign' of some topic, subject, or other object quite different from that implied by its directly perceived characteristics."

Almost any experimental study in which *S*s are required to make perceptual reactions will be found to furnish evidence of a difference in *S*s' observing sets which is correlated with a difference in reactions, and which can be interpreted as an example of this tendency to observe on various levels of abstraction. In the study of constancy phenomena, for example, the exact nature of the set being employed by *S* is always a source of confusion. In equating two color discs, one in bright light, the other in shadow, is *S* comparing two *sensations* of brightness, or is he comparing two bright *objects?* The crucial effect of attitude in determining such measurements is stressed by Woodworth (**22**). And it is apparent from his account that on occasion *S* may assume a concrete attitude (". . . 'critical' attitude, like that of a painter who tries to recapture the stimulus colors on his canvas") and exhibit a low degree of constancy; on other occasions the *S* may assume a more abstract attitude (". . . the naive attitude of everyday observation, which undoubtedly is the attitude of looking for object colors") and exhibit a high degree of constancy.

This same difference in level of abstraction can be found to operate in studies of the influence of stereotypes in perception. The typical case of the influence of a stereotype is one in which *S* has reacted on a relatively high level of abstraction to a particular stimulus which bears a certain label. Thus the average white American characteristically reacts to any object labelled 'negro' as a category well stocked with fixed ideas concerning the object's qualities. It is relatively easy for example to catch one of his guard with the question, "What color is So-and-so?," and receive the answer, "Why, black," although he probably never saw a black person in real life. The operation of stereotypes in experimental studies is usually demonstrated as a trend in *average* judgments. Seldom does every *S* in a group show the effect of the stereotype. Occasionally, as in Sherif's study (**18**), the reason for failures of the stereotype to work is investigated, and it is found that the atypical *S*s attended to the stimulus-in-itself, i.e., they refused to react on the abstract level of the stereotype.

The problem of the stimulus error has always been a source of confusion and misconception among psychologists. From the very beginning of experimental psychology it was recognized that *S*s observe under two different sets. One of these was described as a passive, sensorial set; the other, as an active, critical, common sense, meaning-identifying set. Titchener maintained that the former was the only one appropriate to psychological investigations, and that, hence, observing under the second attitude was an error, the stimulus error. The nature of this error has been difficult to define or grasp clearly only in theoretical discussions of it; in the context of an actual experiment it has always been obvious (*cf.*, for example, Fernberger (**6**) and George (**7**)). The sources of confusion about the stimulus error have probably been threefold, and all of them are removed when observing sets are interpreted in terms of levels of abstraction. One source of confusion has been in the dual meaning of the term 'abstraction.' It is frequently said, in descriptions of trained introspection, that when the *S* takes the content attitude he is required to 'abstract' for attending to one particular feature of the stimulus, such as loudness or saturation. As employed in this paper 'abstraction' means almost the opposite. It refers to consideration of the stimulus as a whole and as a meaningful object. A second source of misconception has been the quite general failure to recognize that this diversity of observing sets is not peculiar to introspective experimentation, but is a very common occurrence of everyday life. Different people at the same time, and the same person at different times, observe under the guidance of sets of different levels of abstraction. The third source of confusion, probably the most serious, is the failure to recognize that there are more than just two observing attitudes, that there is really an indeterminant number of them—one corresponding to every possible level of an abstraction hierarchy. One may observe a blue area of certain relative brightness, or he may observe a blue bottle, he may observe an ink-bottle, he may observe a tool used in human communication—and so on from species to genus to ever higher levels of abstraction. It should be noticed, in addition, that there is obvious variation in *direction* a set may take, as it moves to higher levels. For instance, the set to observe ink-blots may not only shift from a particular wolf, to the class of wolves, to mammals, to animals, to animate objects (from lower to higher levels of abstraction), but, as has been frequently demonstrated in experiment, the set may shift, in direction, under the determination of instructions to a particular spider, to the class of spiders, etc. Whether or not failure to limit the *S*s' sets to the lowest possible level of concreteness (the stimulus error) in psychophysical experiments is an error may be a matter of opinion, but failure to take cognizance

of the fact that *S*s vary widely in their level of abstractness and to attempt to control them is certainly an error in any field of experimentation in which *S*s are called upon to make perceptual reactions.

For an exhaustive exposition of the practical and theoretical implications of the perceptual levels of abstraction the reader should be referred to Korzybski's *Science and Sanity* (**12**). That is the theme of his entire book. Although there are no references made in it to the experimental psychology of perception, the psychologist can hardly leave reading it without the conviction that the abstractness of observing set is an extremely important variable in the determination of human behavior.

(Manuscript received October 20, 1943)

References

1. Aveling, F. The relation of thought-process and percept in perception. *Brit. J. Psychol.*, 1911, **4**, 211–227.
2. Bullough, E. The 'perceptive problem' in the aesthetic appreciation of single colours. *Brit. J. Psychol.*, 1908, **2**, 406–463.
3. Dewar, H. A comparison of tests of artistic appreciation. *Brit. J. educ. Psychol.*, 1938, **8**, 29–49.
4. Eysenck, H. J. The general factor in aesthetic judgments. *Brit. J. Psychol.*, 1940, **31**, 94–102.
5. Feasey, L. Some experiments on aesthetics. *Brit. J. Psychol.*, 1921, **12**, 253–272.
6. Fernberger, S. W. An experimental study of the 'stimulus error.' *J. exp. Psychol.*, 1921, **4**, 63–76.
7. George, S. S. Attitude in relation to the psychophysical judgment. *Amer. J. Psychol.*, 1917, **28**, 1–37.
8. Goldstein, K. The meaning of words. *J. Psychol.*, 1936, **2**, 301–316.
9. Granit, A. R. A study on the perception of form. *Brit. J. Psychol.*, 1921, **12**, 223–247.
10. Guilford, J. P. *Fundamental statistics in psychology and education.* New York: McGraw-Hill, 1942. P. 169.
11. Harris, A. J. Affective contrast between modalities. *Amer. J. Psychol.*, 1932, **44**, 289–299.
12. Korzybski, A. *Science and sanity.* New York: Science Press, 1933.
13. Nakashima, T. Contributions to study of the affective processes. *Amer. J. Psychol.*, 1909, **20**, 157–193.
14. Neff, W. S. Perceiving and symbolizing: an experimental study. *Amer. J. Psychol.*, 1937, **49**, 376–418.
15. Peters, H. N. Experimental studies of the judgmental theory of feeling: IV. Retention of the effects of learned reactions on affective judgments. *J. exp. Psychol.*, 1939, **24**, 111–134.
16. Peters, H. N. Experimental studies of the judgmental theory of feeling: V. The influence of set upon the affective values of colors. *J. exp. Psychol.*, 1943, **33**, 285–298.
17. Pratt, C. C. Some qualitative aspects of bi-tonal complexes. *Amer. J. Psychol.*, 1921, **32**, 490–515.
18. Sherif, M. An experimental study of stereotypes. *J. abnorm. (soc.) Psychol.*, 1935, **29**, 371–375.
19. Valentine, C. W. The method of comparison in experiments with musical intervals and the effects of practice on the appreciation of discords. *Brit. J. Psychol.*, 1914, **7**, 118–135.
20. Wells, E. F. The effect of attitude upon feeling. *Amer. J. Psychol.*, 1930, **42**, 573–580.
21. Williams, E. D., Winter, L., & Woods, J. M. Tests of literary appreciation. *Brit. J. educ. Psychol.*, 1938, **8**, 265–284.
22. Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938. Pp. 607, 620–621.
23. Yokoyama, M. Affective tendency as conditioned by color and form. *Amer. J. Psychol.*, 1921, **32**, 81–107.

PROBLEM SOLUTION BY MONKEYS FOLLOWING BILATERAL REMOVAL OF THE PREFRONTAL AREAS. IV. RESPONSES TO STIMULI HAVING MULTIPLE SIGN VALUES

BY H. F. HARLOW AND T. SPAET*

University of Wisconsin

I. Introduction

The purpose of the following investigation was to study the performance of rhesus monkeys, following bilateral extirpation of the prefrontal areas, in the solution of problems in which a single stimulus acquired multiple sign values. These problems are so designed as to produce an increasing complexity of factors determining the response to the primary stimulus and as such are tests of the performance of monkeys to stimuli which come to elicit varying and antagonistic responses. The design of the behavioral tests and the learning of these tests by four unoperated rhesus monkeys have been described in an earlier publication (2).

II. Subjects

Two rhesus monkeys were used in this study. The prefrontal association areas had been removed in both animals in a one stage operation. Both *Ss* were thoroughly experienced test animals, having been run on discrimination, discrimination-reversal, matching-from-sample and delayed reaction matching-from-sample, and initiation of behavior problems. Details of the operation and the earlier tests have already been reported (3, 7, 4).

III. Apparatus

The monkeys were tested in an apparatus similar to that commonly used at the University of Wisconsin and described in one of the previous studies (7).

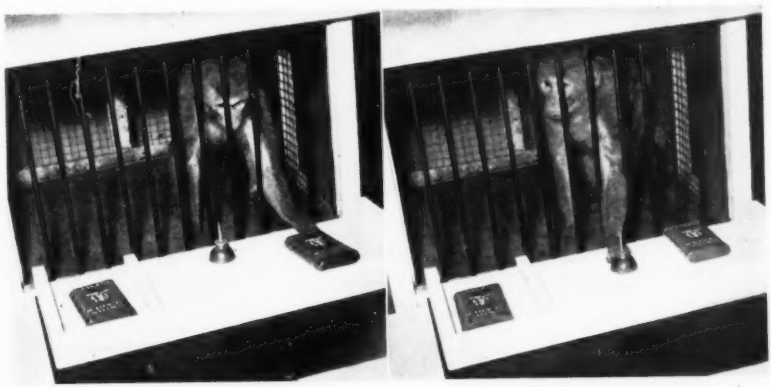
The test tray used in this investigation was divided into two areas. The smaller of these contained the food well over which the sample-object was placed. The larger area contained two food wells over which the two choice-objects were placed. The stimuli used in these tests were a pair of Union Leader tobacco cans and a pair of small copper oil cans.

IV. Procedure

The procedure was very similar to that already described in a problem in which unoperated rhesus monkeys were trained to respond to stimuli having multiple sign values (2), and the reader is referred to this article for procedure details.

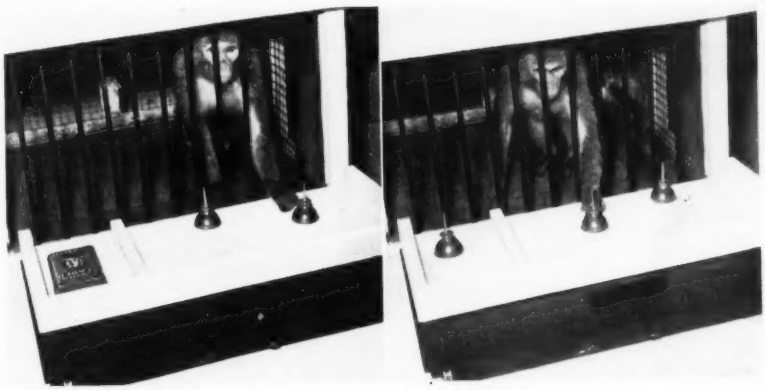
* This work was supported in part by a grant from the Special Research Fund of the University of Wisconsin for 1941–42.

General Test Procedures.—All testing was carried out using a black matching-from-sample tray. Each test trial was initiated by lowering an opaque screen directly in front of the animal's cage and placing food under the appropriate stimulus-object or stimulus-objects. Although the test tray was never moved during testing, it was so placed that each *S* had to reach almost to full arm's length to displace the stimulus objects. A door with a one-way vision screen immediately in front of *E* was then closed, the forward opaque screen was raised, and *S* was allowed to respond. In case of error, the tray was withdrawn, the unattained reward exposed, and the trial rerun.



(*a*) Matching-from-sample
Food under sample

(*b*) Nonmatching-from-sample
No food under sample



(*c*) SDAPH
Food under sample
Right position object arbitrarily correct

(*d*) RSDAH
Food under sample
Left-position object arbitrarily correct

FIG. 1. Responses by animal 54 illustrating four of the sixteen different configurations of the multiple sign problem.

Note: In making the pictures the animal was moved to a different room and a white sliding tray was substituted for the black tray. These changes were made solely for photographic purposes. Photographs of the subject 54 in the actual test situation are available.

Specific Test Procedures.—A stepwise training procedure was used throughout this experiment. The training procedure was designed, in a general way, to necessitate progressively more complex and varied responses to a single stimulus if problem solution was to be attained. Correct responses to representative members of the various tests are illustrated in Fig. I.

Test I: *Alternation of matching-from-sample and nonmatching-from-sample.*

In the matching-from-sample tests, the *S* was required to displace the sample-object and to then choose the identical choice-object. Both sample-object choice and correct choice-object choice were rewarded with food.

To solve the nonmatching-from-sample tests, the *S* had to displace the sample-object, *which was not rewarded with food,* and to then choose the non-identical choice-object. *Solution of the two antagonistic problems was made on the basis of differences in foodsign.* Food was under the sample in the matching tests and was not under the sample in the non-matching tests. The following three steps were used in training for Test I:

Step Ia: Each day's run consisted of 25 matching trials followed by 25 nonmatching trials.

Step Ib: Alternate blocks of five matching and five nonmatching trials were presented each day, a total of 50 trials being given.

Step Ic: The matching and nonmatching trials were presented in an irregular but predetermined order with not more than three matching or three non-matching trials ever succeeding each other.

The criterion for all steps of Test I was the same. The monkey was required to make 45 correct out of 50 trials successively, with the added provision that not more than three errors be made on either matching or nonmatching.

Test II: *Sign-differentiated-antagonistic-position-habits.*

In the sign-differentiated-antagonistic-position-habits the two choice-objects were identical and the sample-object did not match either choice-object.

Food under the sample-object indicated that the most distant (right-position) choice-object was correct. No food under the sample-object was the cue for correctness of the nearest (left-position) choice object. These relationships held regardless of whether tobacco can or oil can was correct. Presence or absence of food under the sample was the only differentiating sign for correctness of the right or left position.

Twenty-five right-position choices and 25 left-position choices were presented each day in an irregular but predetermined order. The criterion for solution was 45 correct trials in a single day's run with the added provision that not more than three errors be made on either the right or the left position.

Test III: *Alternation of matching-from-sample, nonmatching-from-sample, and sign-differentiated-antagonistic-position habits.*

Test III was a combination of the matching and non-matching problem and the SPADH (sign-differentiated-antagonistic-position-habits) problem. This test, like Test I, was run using a three-step procedure.

Step IIIa: Ten matching trials were followed by 10 nonmatching trials and these were followed by 20 SDAPH trials, correct right and left positions being equally represented.

Step IIIb: Blocks of five matching and five nonmatching trials were presented in sequence until a total of twenty trials were completed. Twenty SDAPH trials were then run, correct left and right position responses following in irregular order.

Step IIIc: Ten matching, nonmatching, right-position and left-position RSDAPH trials were presented in irregular, preëstablished order. No particular problem was presented more than three times in succession.

The criterion for test solution was made in terms of the last 100 responses (25 trials of each problem). Forty-five out of 50 correct responses on each of the paired problems of matching and nonmatching, and right and left-position choices of the SDAPH were demanded, with the added provision that not less than 22 correct out of 25 choices be made on each of the four principles involved.

Test IV: *Reversed-sign-differentiated-antagonistic-position-habits.*

Test IV, the reversed-sign-differentiated-antagonistic-position-habits test (designated as RSDAPH), was always arranged with three identical objects, three tobacco cans or three oil cans, on the test tray. The sign values of the food signs in this test were the opposite of the values assigned to them in the SDAPH test. Thus, food under the sample in the RSDAPH test indicated that the left position was correct, and no food under the sample indicated that the right position was correct. All other procedures were the same as those for Test II.

At the end of each day's run of 25 RSDAPH trials, 10 SDAPH trials were presented and were described as maintenance trials.

Test V: *Alternation every five trials of SDAPH and RSDAPH.*

Combination of the SDAPH and RSDAPH was first accomplished by running groups of five trials of each test alternately until a total of 50 trials was completed each day. The criterion for solution was 45 correct out of 50 responses and not more than three errors on either test was permitted.

Test VI: *Random alternation of SDAPH and RSDAPH.*

This test was the same as Test V save for the fact that the two problems were presented in an irregular sequence instead of in alternate blocks of five trials each.

Test VII: *Alternation of matching, non-matching, SDAPH and RSDAPH.*

In this test the matching, nonmatching, SDAPH and RSDAPH problems were presented in irregular sequence. Forty trials a day were run, 10 trials on each problem. The criteria for solution were made in terms of the last hundred trials run. The subjects were required to make 90 out of 100 correct responses and to make not less than 22 correct out of the last 25 responses on each of the four problems.

V. Results

Test I: *Alternation of matching-from-sample and non-matching-from-sample.*

The results for Test I are presented in Table I. The data for this test cannot be directly compared with those of the same test for normal animals, since differences in previous test history existed and since a different stepwise training procedure was used than with the normal monkeys. However, subject 54 did take more trials to solve this problem than did any of the normal monkeys who required from 400 to 1450 trials and subject 55 did take over twice as many trials.

TABLE I

Number of Training Trials to Solve Test I

Subject Number	Test Ia	Test Ib	Test Ic	Total Trials Test Iabc
54	1100	400	150	1650
55	2700	200	250	3150

Test II: *Sign-differentiated-antagonistic-position-habits.*

The data on the SDAPH problem are presented in Table II. Animal 55 required almost three times as many trials to learn this test as the poorest of the normal monkeys. The score made by animal 54 is within the normal range: from 100 to 250 trials.

TABLE II

Number of Training Trials and Number of Correct Responses in the Last 25 Left-position and Right-position Trials for Test II

Subject Number	Training Trials	Left-position Trials	Right-position Trials
54	200	22	23
55	700	23	24

Test III: *Alternation of matching-from-sample, non-matching-from-sample, and sign-differentiated-antagonistic-position-habits.*

Both operated monkeys did poorly on the combined problems of Test III. The four normal animals solved this problem in 100 to 400 trials. The operated monkeys, 54 and 55, required 700 and 1350 trials respectively. The data for the operated animals are presented in Table III.

TABLE III

NUMBER OF TRAINING TRIALS AND NUMBER OF CORRECT RESPONSES IN THE LAST 25 MATCHING, NON-MATCHING AND RIGHT- AND LEFT-POSITION SDAPH TRIALS FOR SOLUTION OF TEST III

Subject Number	Training Trials	Number of Matching Trials	Correct Responses in Last 100 Trials		
			Non-matching Trials	Left-position Trials	Right-position Trials
54	700	23	22	24	25
55	1350	24	22	23	23

Test IV: *Reversed-sign-differentiated-antagonistic-position-habits.*

Subject 54 solved the RSDAPH problem within the range of the number of trials taken by the normal monkeys. Subject 55 failed to solve this problem within 1000 trials, two and a half times the maximum number of trials taken by any unoperated monkey. The results are given in Table IV.

TABLE IV

NUMBER OF TRAINING TRIALS AND NUMBER OF CORRECT RESPONSES IN THE LAST 25 SDAPH AND RSDAPH TRIALS

Subject Number	Training Trials	Number of Correct Responses in Last Test Day	
		SDAPH Trials	RSDAPH Trials
54	300	23	22
55	1000	21*	20*

* Failed to attain criterion for problem solution. It may be noted, nevertheless, that these scores are significantly better than chance at the one percent confidence level.

Test V: *Alternation every five trials of sign-discriminated-antagonistic-position-habits and reversed-sign-discriminated-antagonistic-position-habits.*

Subject 54 was the only one of the operated animals tested on Test V, and she solved it in 1100 trials. The range for the normal animals was 100 to 400 trials.

Test VI: *Random alternation of sign-differentiated-antagonistic-*

position-habits with reversed-sign-discriminated-antagonistic-position-habits.

This problem was solved by subject 54 in 700 trials. The range for the normal animals was from 50 to 350 trials. As in the case of Test V, these data strongly suggest that some loss had occurred as a result of the operation.

Test VII: *Alternation of matching-from-sample, non-matching-from-sample, sign differentiated-antagonistic-position-habits and reversed-sign-differentiated-antagonistic-position-habits.*

Subject 54 failed to solve Test VII, whereas all four normal subjects solved this problem. The range for the normal animals for training trials to effect solution was 100 to 600 trials. Subject 54 failed after 2000 trials which had been set as an a priori limit. In trials 1900 to 2000, she made 22, 18, 22 and 14 correct responses in the 25 trials of matching, non-matching, SDAPH and RSDAPH respectively. A total of 320 additional trials were run without any evidence of improvement of score and the experiment was then discontinued. At one time or another subject 54 had made scores three SD or more beyond chance for 25 consecutive trials of all the problems save the RSDAPH.

VI. Discussion

The scores of the two prefrontal monkeys on the various multiple sign problems described in the paper are consistently inferior to those obtained earlier on 4 normal rhesus monkeys. These data suggest that the extirpation impaired the ability of the operates to make varying and antagonistic responses to a single stimulus depending upon the particular patterning of the total situation. Since the problem was completed in less than two years after the operation, caution should be exercised in assuming that the operation produced permanent loss. It is entirely possible that two years is too short a time to permit complete post-operative recovery.

Even though some deficit appears to exist, the residual abilities of the operated *S*s are striking and represent achievements never previously reported for monkeys following excision of the frontal areas and seldom attained by any subhuman animals. In the combined matching and nonmatching tests the monkeys solved third-order generalizations (**2**) or third-order contingency problems—a creditable performance for any infrahuman organism. The data of this paper indicate that a considerable degree of intellectual sparing followed even the radical operation reported. Similar general results have been described for human patients by Hebb and Penfield (**5**) and by Nichols and Hunt (**6**). In both studies the patients were given over a year to recover from the operations.

The ability of animals, human and subhuman, normal and pathological, on any test, is as dependent upon experimental procedure as it is upon any difficulty intrinsic to the task itself.

The success of the monkeys on the complicated tests of this paper is in large part a result of the nature of the procedure and the long period of adaptation to the test situation. The methods used in this experiment, in their own way, give instructions by the specific cues or signs standing for more abstract generalizations. The use of stepwise procedures may simplify the apparent complexity of the final tasks. To begin with, a single stimulus elicited few differential responses and additional sign-values were not added until the earlier ones had been thoroughly established.

The importance of method in making possible the solution of complex problems by prefrontal animals is suggested in the report of Nichols and Hunt (**6**) on the effects of partial bilateral frontal lobectomy on a human patient. Strong evidence was advanced to show that the *S*s' performance was markedly better when given specific instructions by the experimenters as to method of approach to problems, or when told to look for new systems of attack upon the problems. The inability of human *S*s, following serious brain injury, to shift preëstablished sets and to utilize new methods of approach in the solution of tasks has been emphasized by Goldstein (**1**).

It is more than possible that the best way to illustrate the difference between normal and pathological animals, both human and subhuman is by the use of relatively ineffective test techniques (from a pedagogical point of view). Whenever animals or people are thrust directly into a new situation without adequate preparation, divergence in abilities is exaggerated. Learning to read, learning how to do long division, learning to solve algebraical equations are more difficult tasks than subsequent reading, doing more long division and solving further algebraical equations. The pedagogically good training technique seeks to decrease the difficulty offered by the original learning of a particular skill and thus reduce individual variation in these performances.

It seems certain that serious brain injury reduces general learning ability. The ability to find new and appropriate ways to learn new problems appears to suffer particularly. The answer to this may lie in the fact that learning how to learn is an extremely difficult aspect of the process of acquisition. If this is true, teaching methods become of extreme importance and analysis of procedures is as important as analysis of problems in the determination of the effects of cortical lesions.

(Manuscript received November 12, 1943)

References

1. Goldstein, K. The problem of the meaning of words based upon observation of aphasic patients. *J. Psychol.*, 1936, **2**, 301–316.
2. Harlow, H. F. Responses by rhesus monkeys to stimuli having multiple-sign values. In: McNemar, Q. & Merrill, M. A., edit. *Studies in personality.* McGraw-Hill Co., 1942, 105–123.
3. Harlow, H. F. & Dagnon, J. Problem solution by monkeys following bilateral removal of the prefrontal areas. I. The discrimination and discrimination-reversal problems. *J. exp. Psychol.*, 1942, **32**, 351–356.
4. Harlow, H. F. & Johnson, T. Problem solution by monkeys following bilateral removal of the prefrontal areas. III. Test of initiation of behavior. *J. exp. Psychol.*, 1943, **32**, 495–500.
5. Hebb, D. O. & Penfield, W. Human behavior after extensive bilateral removal from the frontal lobes. *Arch. Neurol. Psychiat.*, 1940, **44**, 421–438.
6. Nichols, I. C. & Hunt, J. M. A case of partial bilateral frontal lobectomy. *Amer. J. Psychiat.*, 1940, **96**, 1063–1083.
7. Spaet, T. & Harlow, H. F. Problem solution by monkeys following bilateral removal of the prefrontal areas. II. Delayed reaction problems involving use of the matching-from-sample method. *J. exp. Psychol.*, 1942, **32**, 424–434.

ACCURACY OF RECOGNITION OF SUBLIMINAL AUDITORY STIMULI

BY JANE W. COYNE, H. E. KING, J. ZUBIN AND C. LANDIS

Psychiatric Institute, Columbia University

If three different and clearly distinguishable tones are sounded through the earphone of an audiometer, they may (when heard) be differentiated and called a low, middle or high tone. Using the standardized audiometric procedure, one may accurately establish the auditory threshold for each tone so that above an individually determined point on the intensity scale one always hears and can differentially distinguish the low, middle or high tone. Below this point on the scale there is an area of variable width in which a subject will report that sometimes he hears a tone and that sometimes he hears nothing. Going beyond this area of uncertainty one finds that throughout the remainder of the lower end of the intensity scale the subject consistently reports that he hears nothing. What will happen if, after he has reported that he heard nothing, one asks him to guess whether it was the low, middle or high tone?

Baker (1) reports an experiment along similar lines to that just described. His subjects did guess and name correctly subliminal tones more frequently than the laws of chance would permit. In the realm of visual phenomena Williams (4), Vinacke (3), and King, Landis and Zubin (2) have reported that such better-than-chance guessing, which they call *subliminal perception*, also occurs.

The present experiment is concerned with one major question and two variations of that question. Can subliminal tones be differentially reported with a degree of accuracy above chance expectancy under certain special conditions? And, when subliminal intensities are systematically varied, will the accuracy of guessing increase as the intensity of the tone approaches the area of uncertainty? Further, when a random noise sound screen is used to artificially raise the auditory threshold, will the accuracy of guessing subliminally presented tones be altered in any way?

Experimental Conditions

A Western Electric Company Type 2A audiometer was used as the stimulus source. Three tones, C^1 (256 c.p.s.), C^2 (1024 c.p.s.) and C^3 (4096 c.p.s.), which we shall call low, middle and high, were employed. The *S* was seated comfortably in a sound-deadened room. The earphone was held in place by a headband and placed over the ear which *S* said he preferred to use. At a dis-

tance of four feet in front of *S* was a small signal lamp which was turned on and off simultaneously with the sounding of the tone. The *S* was instructed to report or to guess the tone presented as soon as the lamp went out. The experimenter and the audiometer were outside the experimental room. Other than the signal lamp there was no illumination in *S*'s room. Almost all experimental sessions were conducted at night to minimize extraneous and distracting noises.

Three determinations of threshold were made using a descending series of tonal intensity, and three using an ascending series. From these six determinations the average between the lowest figure obtained in the ascending and in the descending series was obtained and called the minimal threshold. (Threshold, as we will use the term throughout this report, refers to this minimal level of intensity.) Such thresholds were determined for C^1, C^3 and C^5 for each *S*. For subliminal stimulation the audiometer was set at 10 decibels (high intensity), 15 decibels (middle intensity) and 20 decibels (low intensity) below these individually determined thresholds for each tone. Since we had three tones and three levels of intensity there were nine possible combinations of presentation. These nine combinations were presented to each *S* in random order. Seven identical repetitions of each of these nine combinations, randomized to remove the influence of the order of presentation, were given each *S*, making 63 in all. A similar scheme of randomization was employed in the series with the sound screen. This experimental design is shown in Table I.

TABLE I

EXPERIMENTAL DESIGN

(These combinations were presented in a randomized order)

Tones	Intensity			
	High: 10 db. Below Threshold	Middle: 15 db. Below Threshold	Low: 20 db. Below Threshold	Total
Low: 256 cps	7	7	7	21
Middle: 1024 cps	7	7	7	21
High: 4096 cps	7	7	7	21
Total	21	21	21	63

It is theoretically conceivable that the electrical energy from the audiometer which reaches the earphone and there sets up mechanical energy in the form of sound waves which are transmitted to the tympanum of the *S* may be too slight to activate the aural mechanism. This might be true especially when we were sounding tones at 20 decibels below the minimal threshold.

In order to obviate this theoretical point and to conceal such sounds as do occur in quiet conditions, we made use of a sound screen which had a sound level of 67 decibels as measured at the position of *S*'s head. The sound was produced by permitting compressed air to escape at a constant rate and pressure through a pipette opening. We found that this constant "whis-s-h" did not produce beats, harmonics or other aural sound complications with either of the tones C^1, C^3 or C^5.

Ten *S*s were used; 63 guesses made in the quiet series and 63 in the sound screen series giving a total of 1260 responses.

The *S*s were students, physicians and psychologists. All were free of hearing defect. No one was acquainted with the object of the experiment. In preliminary experimentation, *S*s reported difficulty in remembering the qualitative difference between the three tones. ("I cannot remember the difference in sound between low, middle and high." "How can I guess when I can't remember what they sound like?") To overcome this trouble in the major experiments each tone was sounded supraliminally after every set of six guesses.

Results

In the treatment of the data we must take as our point of departure the number of correct responses which one would obtain from the operation of chance alone. For, if the number of correct guesses is no greater than chance would permit, we need not invoke any other factors to explain our observations. If, however, the number of correct 'guesses' exceeds chance expectancy significantly, the hypothesis that chance alone is sufficient to explain our results is untenable and some other alternative hypothesis such as subliminal perception may be advanced.

Group Data.—As Table II shows, the total number of responses obtained from the group of 10 *S*s was 1260 of which 503, or 40 percent,

TABLE II

Number of Experimentally Obtained Correct Responses Distributed According to Pattern of Experimental Design

Tones	Intensity											
	High: 10 db.			Middle: 15 db.			Low: 20 db.			Total		
	Quiet	Noise	Total	Quiet	Noise	Total	Quiet	Noise	Total	Quiet	Noise	Total
Low: 256 cps	24	28	52	28	24	52	36	29	65	88	81	169
Middle: 1024 cps	26	31	57	32	28	60	35	27	62	93	86	179
High: 4096 cps	26	22	48	29	24	53	27	27	54	82	73	155
Total	76	81	157	89	76	165	98	83	181	263	240	503

were correct judgments. By chance alone only $33\frac{1}{3}$ percent correct would be expected, since there were three possible choices for every response. The excess of correct responses is such that it would be expected to occur less than one time in 100 under the chance hypothesis. Accordingly, we must conclude that the accuracy of report achieved by the group as a whole cannot be attributed to chance, and our alternative hypothesis of subliminal perception (or some other still unidentified factor) may be invoked. The above result holds true for both the quiet condition as well as for the sound screen condition, but was considerably more pronounced for the former than for the latter. For the quiet series the level of accuracy was 42 percent while with the sound screen it was 38 percent.

Individual Data.—Having shown by the analysis of the group data that some factor was operating to produce better-than-chance

guessing, we now turn to the examination of the level of accuracy for each *S*. The analysis of this data is shown in Table III. Taking the results of each *S* and combining both his success in the quiet and in the sound screen series we find that two *S*s gave results significantly better than chance, six exceeded chance but not to the point of statistical significance, and two fell slightly short of chance. We have, then, eight out of 10 individuals exceeding chance, an occurrence having a probability of less than 11 in 1024. This indicates that not only the total score for the entire group exceeded chance expectancy, but that the distribution of individual scores behaved likewise.

TABLE III

THE VALUES OF χ^2 FOR THE NUMBER OF CORRECT AND INCORRECT RESPONSES BY INDIVIDUAL SUBJECTS

Subject	Quiet			Noise			Total		
	No. Correct	χ^2	P	No. Correct	χ^2	P	No. Correct	χ^2	P
1	34	12.07	.001	29	4.57	.03	63	15.75	.001
2	39	23.14	.001	27	2.57	.11	66	20.57	.001
3	19	0.29	.60	22	0.07	.79	41	0.04	.82
4	29	4.57	.03	23	0.29	.59	52	3.57	.06
5	25	1.14	.28	25	1.14	.28	50	2.29	.13
6	20	0.07	.79	25	1.14	.28	45	0.32	.57
7	22	0.07	.79	28	3.50	.06	50	2.29	.13
8	23	0.29	.59	17	1.14	.28	40	0.15	.48
9	25	1.14	.28	21	0.00	1.00	46	0.57	.44
10	27	2.57	.11	23	0.29	.59	50	2.29	.13
Expected	21			21			42		

When the results of the experiment under quiet conditions were analyzed with respect to each *S*, it was found that three were significantly better than chance, five were better but not significantly so, and two fell slightly short of chance. With the sound screen one was significantly better, seven better but not significantly so, one just at chance and one somewhat below chance.

One *S* reported on five occasions that he thought that he heard something when a subliminal tone had been sounded. Four of these 'somethings' he named correctly and one incorrectly. Four other *S*s reported one instance each of hearing something when a subliminal stimulus was presented. Two guesses were correct and two incorrect in these four instances. In all of nine such occurrences six were correct and three incorrect. The separation of these nine instances from either the group or individual data did not change the results either to increase or decrease significantly the general accuracy of the guessing.

Influence of Experimental Factors on Accuracy of Guessing.—We systematically varied three conditions in this experiment: the tone, the intensity of the tone and the ground level of sound (noise screen) which surrounded *S*. Differences between individual *S*s might be called a fourth variable. In order to determine the influence of each factor on the accuracy of guessing, the data were statistically treated by the method of analysis of variance. This analysis was first applied to the data of each individual separately and, after finding that the results were essentially alike, they were combined for group analysis. The results indicate that no one of these factors exerted a significant influence on the accuracy of report. Only the differences between individuals are significant when compared with the variance expected by chance alone. The separate analyses of the data obtained under conditions of quiet yielded similar results but, in addition, indicated that there was a significant tendency for accuracy to increase with the intensity of the tone. The analysis of data obtained with the sound screen gave no significant trends and, in this case, even differences between individuals tended to be no greater than chance would provide. We may conclude that the introduction of the sound screen which raised the auditory threshold seems to reduce individual differences and to disturb the trend between increase in accuracy with intensity. But the general level of accuracy is not significantly greater under quiet than under sound screen conditions.

As an additional check, two *S*s (one from the original group and one new *S*) were tested in a series in which the audiometer was turned off for certain randomized interpolated judgments, without the *S*'s knowledge. Keys and lights were operated as in the regular series. Analysis of the results of these blank presentations approximated chance expectancy almost exactly. This is evidence that the experimental procedure, without the audiometrically produced tones, does not furnish the basis for the better-than-chance guessing.

Summary

The purpose of this experiment was to determine whether auditory stimuli presented at intensities of 10, 15 and 20 decibels below the determined minimal threshold can be recognized more often than chance would demand. The three stimuli selected for presentation were tones C^1-256, C^3-1024 and C^5-4096 c.p.s. The minimum intensity threshold for each of these frequencies was determined and, after the *S* became sufficiently acquainted with the tones so that he could recognize their identity in supraliminal presentations, these sounds were presented subliminally in random order. Half of the observations, 63 for each *S*, were taken under very quiet conditions

and half under the condition of noise produced by a sound screen. The results indicate that under both experimental conditions the accuracy of discrimination report is better than chance. The chance expectancy of correct report under the conditions of our experiment was $33\frac{1}{3}$ percent; during quiet conditions the accuracy of report was 42 percent; during the sound screen, 38 percent.

Under quiet condition there was a definite positive relationship between intensity of the presented sound and accuracy. This did not hold true for the sound screen condition. Furthermore, while individual differences were rather prominent under the former, they all but disappeared under the latter.

(Manuscript received October 5, 1943)

References

1. Baker, L. E. The influence of subliminal stimuli upon verbal behavior. *J. exp. Psychol.*, 1937, **20**, 84–100.
2. King, H. E., Landis C. & Zubin, J. Visual subliminal perception where a figure is obscured by the illumination of the ground. *J. exp. Psychol.* (In press)
3. Vinacke, W. E. The discrimination of color and form at levels of illumination below conscious awareness. *Arch. Psychol.*, 1942, **38**, No. 267.
4. Williams, A. C. Perception of subliminal visual stimuli. *J. Psychol.*, 1938, **6**, 187–199.

A SIMPLIFIED OSCILLATOR SUITABLE FOR AUDITORY RESEARCH AND AUDIOMETRY

BY PETER KELLAWAY* AND GILBERT BRIGHOUSE

Occidental College

Most laboratories have frequent use for an inexpensive, easily maintained audio-oscillator which possesses a high degree of accuracy and dependability. The instrument described here offers the following advantages:

1. The frequency range is continuous from 15 to 65,000 cycles per sec. Any frequency within that range is obtainable by the simple adjustment of the band switch and frequency control.
2. It possesses marked stability (less than plus or minus 0.2 percent drift at 100 cycles per sec.).
3. Since it is not of the beat frequency type, the constant adjustment to zero beat is obviated.
4. The output is calibrated in decibels.
5. The cost of construction and maintenance is low.
6. Operation is simple; only three controls are used: band selection switch, frequency selection control, and output voltage control (calibrated in decibels). The four frequency bands cover respectively the ranges 15–150 cps, 150–1200 cps, 1200–10,000 cps, 10,000–65,000 cps.

I. Principle of Operation

The principle of operation is the automatic adjustment of the ratio between the regenerative and degenerative feedback of the output and input circuits. The low thermal drift is maintained through the use of low temperature coefficient elements in the resistance network. The regenerative feedback is fixed in magnitude and is fed back directly to the grid of the 6SJ7 through this resistance capacity network. The effective "Q" of this network is maintained at a high level by the degenerative feedback, assuring frequency stability.

The degenerative feedback is fed through a resistance into two lamps in the cathode circuit of the 6SJ7. These lamps are Mazda type S6 which have a very positive resistance-temperature characteristic; therefore, the amount of feedback voltage drop across the

* Now at McGill University, Montreal.

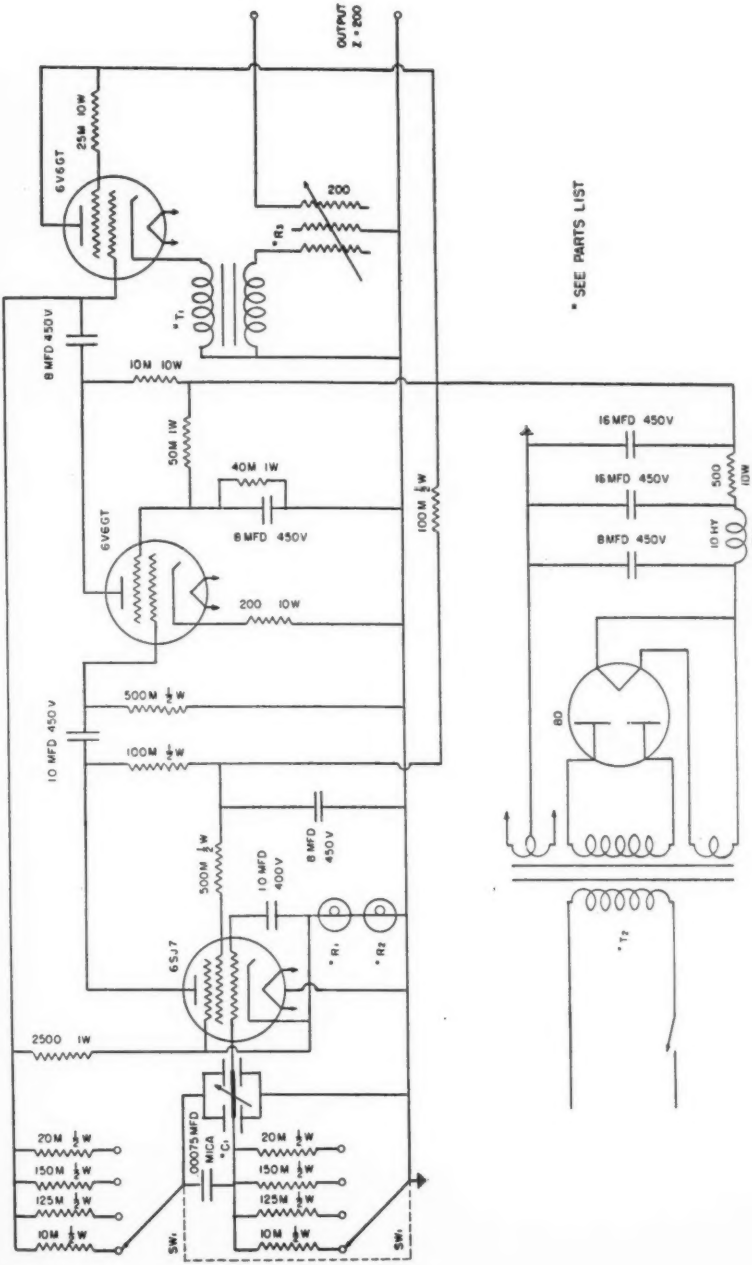


FIG. 1. A simplified oscillator.

lamps is determined by the change in magnitude of the audio current fed back from the first 6V6. The second 6V6 acts as an amplifier, and inspection of the circuit will reveal that it entirely isolates the oscillator from the external circuit; therefore, various output impedances can be used without any essential reaction on the oscillator.

The balance of the capacity network to ground is maintained through the adjustment of the trimmer condensers in parallel with the gang and balancing condensers. In this way any asymmetry of wave form arising from imbalance of the sections of the capacity network to ground is compensated for.

II. Parts List

The output transformer T_1 was designed to match the 6V6 to the 200 ohm line. The attenuator was designed for 200 ohm constant impedance to both sides.

A headphone was made having a 200 ohm impedance at 1000 cycles per second. This was done by experimentation—no values can be given because they vary with the type of material used.

The instrument in use at Occidental College was calibrated by means of a cathode-ray oscilloscope, using the 60 cycle AC line as a frequency standard.

(Manuscript received November 3, 1943)

AMERICAN PSYCHOLOGICAL PERIODICALS

American Journal of Psychology—Ithaca, N. Y.; Cornell University. Subscription $6.50. 624 pages annually. Edited by K. M. Dallenbach, Madison Bentley, and E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

Journal of Genetic Psychology—Provincetown, Mass.; The Journal Press. Subscription $14.00 per annum (2 volumes). 1000 pages annually. Edited by Carl Murchison. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1891.

Psychological Review—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $5.50. 540 pages annually. Edited by Herbert S. Langfeld. Bi-monthly. General psychology. Founded 1894.

Psychological Monographs—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $6.00 per volume. 500 pages. Edited by John F. Dashiell. Without fixed dates, each issue one or more researches. Founded 1895.

Psychological Bulletin—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $7.00. 665 pages annually. Edited by John E. Anderson. Monthly (10 numbers). Psychological literature. Founded 1904.

Archives of Psychology—New York, N. Y.; Columbia University. Subscription $6.00 per volume. 500 pages. Edited by R. S. Woodworth. Without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $5.00. 560 pages annually. Edited by Gordon W. Allport. Quarterly. Founded 1906.

Journal of Educational Psychology—Baltimore, Md.; Warwick & York. Subscription $6.00. 720 pages annually. Edited by J. W. Dunlap, P. M. Symonds, and H. E. Jones. Monthly except June to August. Founded 1910.

Psychoanalytic Review—New York, N. Y.; 64 West 56th St. Subscription $6.00. 500 pages annually. Edited by Smith Ely Jelliffe. Quarterly. Founded 1913.

Journal of Experimental Psychology—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $7.00 per annum. 520 pages annually. Edited by Samuel W. Fernberger. Bi-Monthly. Founded 1916.

Journal of Applied Psychology—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $6.00. 480 pages annually. Edited by Donald G. Paterson. Bi-monthly. Founded 1917.

Journal of Comparative Psychology—Baltimore, Md.; Williams & Wilkins Co. Subscription $14.00 per annum (2 volumes). 1000 pages annually. Edited by Roy M. Dorcus. Bi-monthly. Founded 1921.

Comparative Psychology Monographs—Baltimore, Md.; Williams & Wilkins Co. Subscription $6.00 per volume. 400 pages. Edited by Roy M. Dorcus. Without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Provincetown, Mass.; The Journal Press. Subscription $7.00. 500 pages annually. Edited by Carl Murchison. Bi-monthly. Each number one complete research. Child behavior, animal behavior, and comparative psychology. Founded 1925.

Psychological Abstracts—Northwestern University, Evanston, Illinois; American Psychological Association, Inc. Subscription $7.00. 700 pages annually. Edited by Walter S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

Journal of General Psychology—Provincetown, Mass.; The Journal Press. Subscription $14.00 per annum (2 volumes). 1000 pages annually. Edited by Carl Murchison. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Journal of Social Psychology—Provincetown, Mass.; The Journal Press. Subscription $7.00. 500 pages annually. Edited by John Dewey and Carl Murchison. Quarterly. Political, racial, and differential psychology. Founded 1929.

Psychoanalytic Quarterly—Albany, N. Y.; 372–374 Broadway. Subscription $6.00. 560 pages annually. Edited by Bertram D. Lewin and others. Quarterly. Founded 1932.

Character and Personality—Durham, N. C.; Duke University Press. Subscription $2.00. 360 pages annually. Edited by Karl Zener. Quarterly. Founded 1932.

Journal of Psychology—Provincetown, Mass.; The Journal Press. Subscription $14.00 per annum (2 volumes). 800–1200 pages annually. Edited by Carl Murchison. Quarterly. Founded 1936.

Psychometrika—University of Chicago, Chicago, Ill.; Psychometric Society. Subscription $10.00. 320 pages annually. Edited by L. L. Thurstone and others. Quarterly. Quantitative methods in psychology. Founded 1936.

Psychological Record—Bloomington, Ind.; Principia Press. Subscription $4.00. 500 pages annually. Edited by J. R. Kantor and C. M. Louttit. Without fixed dates, each number a single research. General psychology. Founded 1937.

Journal of Consulting Psychology—Lancaster, Penn.; Science Printing Co. Subscription $3.00. 240 pages annually. Edited by J. P. Symonds. Bi-monthly. Founded 1937.